

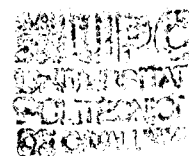


# XARXES NEURONALS PER A LA GENERACIÓ DE DISSENYS EN BLOCS

Tesi Doctoral

Autor: Pau Bofill i Soliguer  
Directora: Carme Torras i Genís





A la Maite, la Mònica i la Mireia. I a la Lourdes, la meva germana petita.

συνακ.





## Resum

Aquest treball descriu l'avaluació de diversos *algorismes* de cerca, basats en xarxes neuronals optimitzadores, aplicats a una família de *problemes*, la generació de dissenys en blocs.

Donat un conjunt  $(v, b, u)$  de paràmetres ( $v$  files,  $b$  columnes i  $u$  uns), un *disseny en blocs* és tota configuració  $v \times b$  binària que compleixi les propietats següents:  $u$  uns,  $r$  uns per fila,  $k$  uns per columna, i correlació  $\lambda$  entre parelles de files. Els valors  $[u, r, k, \lambda]$  s'anomenen aquí *descriptors* del disseny i, donat que han de ser enters, restringeixen l'admissibilitat dels paràmetres. L'admissibilitat, però, no implica l'existència del disseny corresponent.

Un algorisme d'optimització es pot descomposar en una *funció de cost*, que conforma el paisatge, i una *estratègia* de cerca, que determina la manera d'explorar-lo. El treball proposa un repertori de funcions de cost, basades en el nombre de parelles com a mesura de distribució de cada una de les propietats del disseny. L'estructura resultant correspon directament a una xarxa optimitzadora amb connexions d'aritat quatre.

Les estratègies que s'avaluen són tres. Dues estratègies clàssiques (*cerca descendent* i *recuita simulada*), i una estratègia pròpia, la *cerca cooperativa*, que explora l'espai de cerca en paral·lel amb diverses instàncies de xarxa evolucionant en *esquadra*, en la direcció d'energia mitjana decreixent. La reducció progressiva del radi (màxima distància de hamming al centre de l'esquadra), condueix la cerca cap a valls cada cop més profundes.

Donat un cas experimental particular, el cost esperat fins la primera solució és la variable escollida com a *resultat* experimental, ja que té en compte els recursos invertits. Problema, funció i estratègia constitueixen els *factors* principals de l'avaluació experimental. Però, tenint en compte els subfactors que hi intervenen (coeficients de les funcions de cost i paràmetres de les estratègies), una anàlisi factorial completa és intractable. L'experimentació, aleshores, s'ha de planificar en tres etapes: una etapa d'entrenament, per ajustar coeficients i paràmetres, una etapa de comparació, per analitzar les prestacions de cada un dels factors, i una tercera etapa per a l'aplicació del millor algorisme trobat a problemes de tamany creixent.

Els resultats de l'etapa de comparació (sobre els 25 problemes més petits) mostren una forta interacció entre els factors. La recuita simulada (24 problemes resolts) és l'estratègia de cerca que resulta més eficient. La cerca cooperativa supera a la cerca descendent en el nombre de problemes resolts (22 contra 13), però, en els problemes que resolen totes dues, resulta menys eficient. De les 9 estructures proposades, la funció de cost que va millor resulta ser la més senzilla. Els resultats de la tercera etapa mostren que, tot i que alguns problemes són intrínsecament fàcils (o difícils), el nombre de cassos que queden irresolts creix amb el tamany del problema. En conjunt, la generació de dissenys en blocs ha estat un bon banc de proves per als algorismes d'optimització neuronals.

Finalment es proposa una nova família de configuracions combinatòries, que anomenem *dissenys màximament equilibrats*, que admeten fins a dos valors correlatius dels últims tres descriptors, evitant així les restriccions d'admissibilitat. S'analitzen algunes de les seves propietats, i es verifiquen experimentalment. De 38 cassos escollits per a les proves, 26 s'han pogut resoldre. A efectes de generació, doncs, els dissenys màximament equilibrats no són gaire més difícils que els dissenys en blocs.



# **Index general**

<b>1</b>	<b>Antecedents i objectius</b>	<b>1</b>
1.1	El problema: Generació de dissenys en blocs incomplets equilibrats .	1
1.2	L'algorisme de cerca: Xarxes neuronals optimitzadores . . . . .	7
1.2.1	El model discret de Hopfield i la cerca descendent . . . . .	8
1.2.2	La màquina de Boltzmann i la recuita simulada . . . . .	9
1.2.3	Xarxes d'ordre superior . . . . .	12
1.3	Objectius de la tesi i organització de la memòria . . . . .	13
<b>2</b>	<b>Funcions de cost per a la generació de dissenys amb xarxes optimitzadores</b>	<b>17</b>
2.1	Notació i terminologia . . . . .	18
2.2	Mesures de distribució per a les propietats d'un disseny . . . . .	19
2.2.1	Mesures de distribució basades en el nombre de parelles . . . .	19
2.2.2	Mesures de les propietats d'un disseny . . . . .	24
2.3	Optimalitat . . . . .	28
2.3.1	Valors òptims . . . . .	29
2.3.2	Condicions necessàries i suficients d'optimalitat . . . . .	31
2.4	Funcions de cost per a la generació de dissenys . . . . .	33
2.4.1	Definició i forma normal . . . . .	33
2.4.2	Condicions de validesa d'una funció de cost . . . . .	34
2.4.3	Elecció d'un repertori . . . . .	35
2.5	La xarxa optimitzadora . . . . .	36
2.6	Funcions de cost de tipus $H$ . Error quadràtic dels camps locals . . . .	39
<b>3</b>	<b>Estratègies de cerca</b>	<b>42</b>
3.1	L'algorisme d'optimització. Marc general . . . . .	42
3.2	Cerca descendent, CD . . . . .	44
3.3	Recuita Simulada, RS . . . . .	44
3.4	Cerca cooperativa CC, la nova estratègia . . . . .	46
3.4.1	Descripció de l'estratègia . . . . .	46
3.4.2	Formalització en el marc global . . . . .	49

<b>4</b>	<b>Disseny dels experiments per a l'avaluació dels algorismes</b>	<b>50</b>
4.1	Mesures experimentals . . . . .	50
4.1.1	Cost de relaxació . . . . .	50
4.1.2	Caracterització dels resultats d'un cas experimental. Nombre esperat d'intents i cost esperat fins la primera solució . . . . .	51
4.1.3	Estimació dels resultats en un entorn multiprocessador. L'experiment paral·lel . . . . .	53
4.1.4	El cas de la cerca cooperativa. Execució d'S descensos en esquadra . . . . .	54
4.2	Factors experimentals i planificació dels experiments . . . . .	55
<b>5</b>	<b>Descripció dels experiments i anàlisi dels resultats</b>	<b>60</b>
5.1	Primera etapa: Entrenament . . . . .	60
5.1.1	Experiments preliminars. Caracterització de $z$ i $w$ . . . . .	60
5.1.2	Preselecció de les funcions de cost . . . . .	64
5.1.3	Optimització experimentals dels paràmetres de la RS . . . . .	66
5.1.4	Paràmetres estàndard per a la CC . . . . .	68
5.1.5	Altres proves . . . . .	70
5.2	Segona etapa: Comparacions . . . . .	71
5.2.1	Selecció de la millor funció per a cada estratègia . . . . .	72
5.2.2	Comparació entre estratègies . . . . .	72
5.2.3	Validació dels paràmetres estàndard de la RS . . . . .	73
5.2.4	Anàlisi dels resultats en funció del tamany del problema . . . . .	74
5.3	Tercera etapa: Aplicació del millor algorisme a problemes de tamany creixent . . . . .	76
<b>6</b>	<b>Dissenys màximament equilibrats</b>	<b>95</b>
6.1	Definició i antecedents . . . . .	95
6.2	Algunes propietats dels disseny màximament equilibrats . . . . .	96
6.3	Funcions de cost per a la generació de dissenys màximament equilibrats	100
6.4	Generació experimental de dissenys màximament equilibrats . . . . .	102
<b>7</b>	<b>Conclusions</b>	<b>108</b>
7.1	Aportacions i altres treballs relacionats . . . . .	108
7.2	Resultats de l'avaluació experimental . . . . .	109
7.2.1	Aspectes metodològics . . . . .	109
7.2.2	Conclusions dels experiments . . . . .	110
7.3	Discussió . . . . .	112
7.4	Perspectives . . . . .	116

# Capítol 1

## Antecedents i objectius

La *cerca* és el procés d'explorar un *espai* amb un determinat *objectiu* seguint una estratègia. L'anàlisi combinatòria estudia aquells casos en que l'espai està definit implícitament per les combinacions possibles d'un nombre finit d'elements discrets. L'objecte d'estudi, el *problema*, són aquelles *configuracions* o combinacions concretes que compleixen determinades propietats. L'optimització combinatòria estudia els *algorismes* de resolució d'aquests problemes que es basen en la definició d'una *funció de cost* que assigni un valor a cada una de les configuracions, de tal manera que la solució coincideixi amb l'*òptim* (màxim o mínim) de la funció. La manera d'explorar aquest *paisatge* dependrà, aleshores, de l'*estratègia de cerca* que fem servir.

Aquest treball proposa i estudia alguns models de xarxes neuronals optimitzadores aplicades a la generació de dissenys en blocs. El problema, la construcció d'un disseny, consisteix en posar  $u$  peces sobre un tauler de  $v \times b$  files i columnes de tal manera que cada fila tingui  $r$  peces, cada columna en tingui  $k$ , i que les peces de cada parella de files coincideixin en  $\lambda$  posicions. L'espai de cerca són totes les configuracions possibles del tauler. I els algorismes que farem servir es basen en alguns models de xarxa neuronal que incorporen, de manera integrada, funció de cost i estratègia de cerca.

En aquest capítol es presenten els antecedents (disseny i xarxes optimitzadores) i es descriuen els objectius i organització de la memòria.

### 1.1 El problema: Generació de dissenys en blocs incomplets equilibrats

Els dissenys en blocs provenen de l'estadística en l'àmbit del disseny d'experiments [Yates, 35, Fisher, 40]. Suposem, per exemple, que volem comparar  $v$  *varietats* d'una llavor, per veure quina d'elles produeix millor collita. Si disposem d'un territori homogeni prou gran hi plantarem una mostra de cada varietat i estudiarem la mitjana de cada collita. Però si el territori de que disposem no és homogeni, caldrà subdividir-lo en parcel·les que ho siguin. La collita aleshores dependrà de dos

factors: la varietat de llavor que hem fet servir i la fertilitat de la parcel·la on s'ha plantat. L'estratègia habitual, en aquest cas, consisteix en fer *blocs* o subconjunts de les varietats, i plantar cada bloc en una parcel·la diferent. Si cada bloc admet totes les varietats, direm que els blocs són *complets*. En cas contrari, i per  $b$  blocs, el problema de construir un *disseny en blocs incomplets equilibrat* consisteix en decidir quines varietats cal incloure a cada bloc de manera que:

- i) Cada varietat estigui inclosa en el mateix nombre de blocs.
- ii) Tots els blocs incloguin el mateix nombre de varietats.
- iii) Cada parella de varietats coincideixi en el mateix nombre de blocs.

Les condicions i) i ii) defineixen un *disseny en blocs*, i la iii) defineix un *disseny equilibrat*. La condició d'equilibri garanteix que, globalment, la comparació entre parelles de varietats es fa amb la màxima precisió.

Mentre l'estadística s'interessa per l'anàlisi dels resultats de l'experiment, l'anàlisi combinatòria estudia els dissenys en blocs en si, les seves propietats i la seva construcció (vegeu, p.ex., [Anderson, 89], com introducció, o [Hall, 86, Street & Street, 87]). En termes de combinatòria, doncs,

**Definició 1.1** *Un disseny en blocs incomplets equilibrat és una família de  $b$  subconjunts d'un conjunt donat  $S$  amb  $v$  elements tal que, per a valors fixos de  $k$  i  $\lambda$ , amb  $k < v$  i  $\lambda > 0$ ,*

- i) *cada subconjunt té exactament  $k$  elements, i*
- ii) *qualsevol parella d'elements d' $S$  apareix exactament en  $\lambda$  subconjunts.*

Com veurem, de les condicions anteriors es deriva que

- iii) *cada element pertany exactament a  $r$  subconjunts.*

Els valors  $(v, b, r, k, \lambda)$  són els *paràmetres* del disseny, els elements del conjunt  $S$  són les *varietats*, i els subconjunts són els *blocs*. El terme "família" indica que pot haver-hi blocs repetits.

Els dissenys en blocs es poden representar per mitjà de la seva *matriu d'incidència*, que és la terminologia que farem servir al llarg d'aquest treball.

**Definició 1.2** *Sigui  $A \equiv [x_{ij}]$  una configuració donada, sobre l'espai  $\mathcal{A} - v \times b$  de configuracions binàries amb  $v$  files i  $b$  columnes. Sigui  $o = \sum_{i=1}^v \sum_{j=1}^b x_{ij}$  el nombre d'uns de la configuració  $A$ ,  $r_i = \sum_{j=1}^b x_{ij}$  el nombre d'uns de la fila  $i$ ,  $k_j = \sum_{i=1}^v x_{ij}$  el nombre d'uns de la columna  $j$ , i  $\lambda_{ik} = \sum_{j=1}^b x_{ij}x_{kj}$  la correlació entre les files  $i$  i  $k$ <sup>1</sup>. Direm que  $A$  és la matriu d'incidència d'un disseny en blocs amb paràmetres  $(v, b, r, k, \lambda)$  si i només si es compleixen les propietats següents:*

<sup>1</sup>Sempre que es pugui distingir clarament pel context, el símbol  $k$  denotarà, d'una banda, el nombre d'uns d'una columna  $i$ , de l'altra, el subíndex auxiliar d'una fila.

- i) Nombre d'uns just:  $o = u$ , amb  $u = vr = kb$ .
- ii) Files uniformes:  $r_i = r$ ,  $i = 1, \dots, v$ .
- iii) Columnes uniformes:  $k_j = k$ ,  $j = 1, \dots, b$ .
- iv) Equilibri:  $\lambda_{ik} = \lambda$ ,  $i = 1, \dots, v-1, k = i+1, \dots, v$ .

L'entrada  $x_{ij} \in \{0, 1\}$  representa, aleshores, la incidència de l'element  $i$  al bloc  $j$ , cada fila indica a quins blocs pertany un element donat, i cada columna indica quins elements pertanyen al bloc corresponent. Com abans, les condicions i) i ii) es deriven de les altres dues, però les propietats que descriuen ens faran servei al llarg del treball.

A la figura 1.1 es mostren algunes configuracions de l'espai  $\mathcal{A} = 6 \times 10$ . Es sabut que, en aquest espai, tots els dissenys que existeixen tenen per paràmetres  $(6, 10, 5, 3, 2)$ , amb un total de 30 uns. La primera configuració, A1, mostra un exemple de disseny. La segona, A2, il·lustra la noció d'*isomorfisme*. Malgrat que aparentment és ben diferent d'A1, s'ha obtingut per permutació de les seves files i columnes, tal com s'indica al peu. De fet, es sabut que *tots* els dissenys que existeixen amb aquests paràmetres són isomorfs entre si. El següents dos exemples mostren que és possible satisfer la condició d'equilibri amb files i columnes mal repartides, sempre que el nombre d'uns sigui menor que el nombre just. Però, com mostra la configuració A5, files i columnes uniformes no són suficients per satisfer-la. El darrer exemple, amb 30 uns justos, tot i ser molt similar a A1 no satisfà cap de les altres condicions.

Els paràmetres  $(v, b, r, k, \lambda)$  d'un disseny estan lligats per les dues condicions d'*admissibilitat* següents:

$$vr = bk \quad (1.1)$$

$$r(k-1) = \lambda(v-1), \quad (1.2)$$

que expressen relacions de consistència. El nombre total d'uns és el mateix, tant si els comptem per files com per columnes. I els uns que correlen amb els d'una fila  $i$  donada es poden comptar de dues maneres (figura 1.2): considerant el nombre de columnes que intervenen ( $r$  columnes, amb  $k-1$  uns cada una), o considerant les correlacions de la fila  $i$  amb la resta ( $v-1$  vegades  $\lambda$ ).

Si apliquem el raonament anterior a una fila  $i$  *qualsevol*, donat que en un disseny  $k$  i  $\lambda$  són fixos (condicions i) i ii) de la definició 1.1), tenim que

$$r_i = \frac{\lambda(v-1)}{(k-1)} = r, \quad (1.3)$$

tal com s'expressa a la condició iii).

En aquest treball farem servir  $(v, b, u)$ , amb  $u = vr = bk$ , com a paràmetres *independents*, i el disseny corresponent es denotarà com  $DB(v, b, u)$ . Les propietats de

A1, un disseny (6,10,5,3,2)

1		1	1	1	1	5	2	2	2	2	2
	1	1		1	1	5	2	2	2	2	
1	1	1		1	1	5	2	2	2		
	1	1	1	1	1	5	2	2			
		1	1	1	1	5	2				
1		1	1	1	1	5	2				
3	3	3	3	3	3	3	3	3	3	3	30

a)

A2, un disseny isomorf

1		1	1	1	1	5	2	2	2	2	2
	1		1	1	1	5	2	2	2	2	
1	1		1	1	1	5	2	2	2		
	1	1		1	1	5	2	2			
		1	1	1	1	5	2				
1		1	1	1	1	5	2				
3	3	3	3	3	3	3	3	3	3	3	30

b)

A3, equilibrada

1	1					2	2	2	2	2	2
1	1					2	2	2	2	2	
1	1					2	2	2	2		
1	1					2	2	2			
1	1					2	2				
1	1					2					
6	6	0	0	0	0	0	0	0	0	0	12

c)

A4, equilibrada

1	1	1				3	2	2	2	2	2
1	1	1			1	4	2	2	2	2	
1	1	1		1	1	4	2	2	2		
1	1	1		1	1	4	2	2			
1	1	1	1	1	1	5	2				
1	1	1	1	1	1	5					
6	4	3	2	2	2	2	2	2	2	0	25

d)

A5, files i columnes uniformes

1		1	1	1	1	5	2	2	2	2	2
	1	1	1	1	1	5	1	3	2	2	
1	1	1	1	1	1	5	2	2	2		
	1	1	1	1	1	5	2	1			
		1	1	1	1	5	2				
1		1	1	1	1	5					
3	3	3	3	3	3	3	3	3	3	3	30

e)

A6, 30 uns

1		1	1	1	1	5	2	2	1	3	2
	1	1	1	1	1	5	2	2	3	2	
1	1	1	1	1	1	5	2	2	2		
	1	1	1	1	1	4	1	2			
		1	1	1	1	6	3				
1		1	1	1	1	5					
3	3	3	3	2	4	3	3	3	3	3	30

f)

Figura 1.1: Exemples de configuracions-6×10. Matriu d'incidència (només es representen els 1s), uns per fila, uns per columna, uns totals i matriu triangular de correlacions entre files (llegida en horitzontal, correlació entre les files 1-2, 1-3,...,5-6). a) Un disseny en blocs amb paràmetres (6,10,5,3,2). b) Un disseny isomorf a l'anterior, obtingut per permutació de les files 1-2, 3-4, 5-6, i les columnes 1-2, 3-4, 5-6, 7-8 i 9-10. c) Una configuració equilibrada, però amb files i columnes mal repartides. d) Millor distribució, però files i columnes encara no són uniformes. e) Una configuració amb files i columnes uniformes, que no compleix la condició d'equilibri. f) Una configuració amb el nombre d'uns just, que sembla ben distribuïda però no compleix cap de les altres condicions.



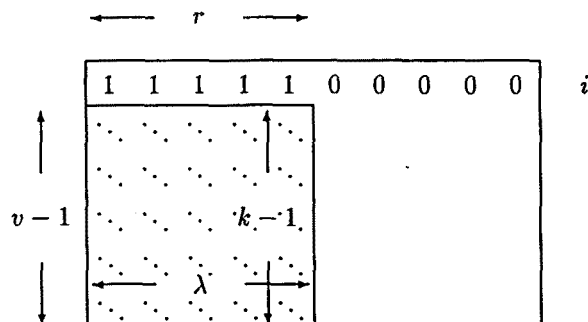


Figura 1.2: Donat un disseny en blocs, i permutant files i columnes per tal d'agrupar els uns de la fila  $i$  de la manera que es mostra, el nombre d'uns dins l'àrea ombrejada es pot comptar de dues maneres. Donat que cada columna té exactament  $k$  uns, les  $r$  columnes implicades en tenen  $k-1$  a l'àrea ombrejada. I, donat que totes les correlacions valen  $\lambda$  i tots els uns de la fila  $i$  estan per sobre de la zona afectada, les  $v-1$  files restants tenen exactament  $\lambda$  uns en aquesta àrea.

la definició 1.2, aleshores, queden resumides pels valors  $[u, r, k, \lambda]$ , que anomenarem *descriptors* del disseny, amb

$$r = u/v \quad (1.4)$$

$$k = u/b, \quad (1.5)$$

$$\lambda = \frac{r(k-1)}{v-1} = \frac{u(u-b)}{bv(v-1)}. \quad (1.6)$$

En aquests termes, els paràmetres  $(v, b, u)$  són admissibles *si i només si* les tres equacions anteriors es satisfan per a valors *enters* d' $r$ ,  $k$  i  $\lambda$ . La relaxació d'aquesta restricció serà objecte d'estudi en el capítol 6.

Citem ara algunes definicions i propietats dels dissenys que ens faran servei més endavant. La seva formulació i demostracions es poden trobar, per exemple, a [Street & Street, 87].

Definicions:

- Un disseny compleix la condició de *lligam* quan la correlació entre cada parella de *columnes* és exactament  $\mu$ ,

$$\mu_{jl} = \sum_{i=1}^v x_{ij}x_{il} = \mu, \quad j = 1, \dots, b-1, \quad l = j+1, \dots, b. \quad (1.7)$$

- La configuració *dual* d'una configuració donada és la resultant d'intercanviar files per columnes.
- Direm que una configuració és *quadrada* quan  $v = b$ .
- La configuració *complementària* d'una configuració donada és la resultant d'intercanviar 1's per 0's.

Propietats:

- Per a qualsevol disseny en blocs,  $v \leq b$  (teorema de Fisher).
- El dual d'un disseny en blocs és un disseny en blocs *si i només si* es tracta d'un disseny quadrat.
- Un disseny en blocs compleix la condició de lligam *si i només si* es tracta d'un disseny quadrat.
- El complementari d'un disseny en blocs és un disseny en blocs.

L'admissibilitat dels paràmetres és condició necessària però no suficient per a l'existència d'un disseny en blocs. Per exemple, per a  $(15, 21, 105)$  els paràmetres són admissibles (amb descriptors  $[105, 7, 5, 2]$ ), però està demostrat que el disseny no existeix (no hi ha cap configuració  $15 \times 21$  que compleixi les condicions). Algunes famílies de dissenys (dissenys que els seus paràmetres satisfan determinades propietats) es poden construir analíticament, per mètodes directes o recursius (vegeu p.ex. [Hall, 86], capítol 15). D'altres s'han resolt amb tècniques algorísmiques (vegeu, p.ex., [Gibbons et al., 77, Corneil & Mathon, 78]). En general, però, l'existència o no existència d'un disseny és un problema obert, que depèn de cada cas particular. Donat un conjunt de paràmetres admissible, el nombre  $N_s$  de solucions no isomorfes a vegades és conegut (o acotat) i a vegades no, però el coneixement (analític) d' $N_s$  no implica, necessàriament, que el disseny s'hagi pogut construir. La situació es troba resumida a [Mathon & Rosa, 90]<sup>2</sup>, on es presenta el llistat de tots els conjunts de paràmetres admissibles amb  $r \leq 41$  i  $3 \leq k \leq v/2$  (els casos amb  $k \leq 2$  són trivials, i els casos amb  $k > v/2$  queden representats pel seu corresponent complementari). Els conjunts de paràmetres es presenten numerats  $(m_i)$  per ordre d' $r$  creixent, i es presenten les cotes conegudes d' $N_s$ , amb les referències bibliogràfiques que les estableixen. En concret,  $N_s = 0$  indica que el disseny no existeix, i  $N_s = ?$  denota un problema obert. La taula 4.1, al capítol 4, en presenta un subconjunt.

Des de la perspectiva de la generació, les solucions analítiques apliquen a famílies infinites, que inclouen dissenys de tamany  $vb \rightarrow \infty$ , però no serveixen per a la resta de casos. Mentre que les tècniques algorísmiques, tot i ser generals, tenen un abast limitat pel tamany del problema. El problema obert més petit que es coneix és  $(22, 33, 264)$  [McKay & Radziszowski, 96], amb  $vb = 726$ . La

<sup>2</sup>La darrera actualització d'aquestes taules es troba a [Mathon & Rosa, 96].

cerca exhaustiva, doncs, no pot encara resoldre casos d'aquestes dimensions. En el cas general, la generació algorísmica de dissenys en blocs és un problema NP-complet [Corneil & Mathon, 78], però a la pràctica, com veurem, la dificultat de generar un disseny depèn més de cada cas particular que no pas del seu tamany. A [Gibbons, 96] es pot trobar una revisió recent de l'estat de la qüestió que, malhauradament, no inclou el llistat sistemàtic de quines tècniques són bones per resoldre quins dissenys.

En aquest treball, l'èmfasi es posa en l'ús dels dissenys en blocs com a *banc de proves* per a la comparació experimental d'algorismes, més que en l'intent de resoldre problemes oberts.

## 1.2 L'algorisme de cerca: Xarxes neuronals optimitzadores

Les xarxes neuronals artificials neixen com a paradigma de computació cap als anys 50 amb els treballs de McCulloch i Pitts, que proposen un model del comportament elèctric de les neurones biològiques, apte per a la computació [McCulloch & Pitts, 43]. I, després d'un període de latència, reapareixen amb força a principis dels anys 80, amb l'aparició d'una sèrie de treballs que obren noves perspectives, com [Hopfield, 82, Rumelhart et al., 86, LeCun, 85, Hinton et al., 84, Kohonen, 87, Grossberg, 87, Barto et al., 81, Jordan, 86, Arbib, 89], entre altres. Els models i aplicacions de xarxes neuronals que s'han proposat són molt diversos (com a introducció podeu consultar, p. ex., [Bofill, 92, Torras, 89, López et al., 89, Lippmann, 87], o bé manuals més extensos com, p.ex., [Rumelhart & McClelland, 86, Kohonen, 87, Hertz et al., 93]).

El denominador comú de tots aquests models és l'existència d'un nombre important d'*unitats* senzilles de procés (les neurones), fortament interconnectades entre si per mitjà de *connexions* ponderades (les sinapsis), que calculen el seu estat d'activació segons una *regla local* del tipus

$$\phi_i = \sum_{j \neq i} w_{ij} x_j - \theta_i, \quad (1.8)$$

$$x_i \leftarrow f(\phi_i) \quad (1.9)$$

on  $\phi_i$  representa l'entrada neta o *camp local* de la unitat  $i$ ,  $x_i$  representa el seu estat d'activació,  $w_{ij}$  és el *pes* de la connexió entre les unitats  $i$  i  $j$ ,  $\theta_i$  és el *llindar d'activació* de la unitat  $i$ , i  $f$  és algun tipus de funció no lineal que acota el marge de valors permesos a  $x_i$ . En la majoria d'aplicacions els pesos  $w_{ij}$  de les connexions i els llindars  $\theta_i$  de les unitats s'entrenen a partir d'exemples i es diu que les xarxes *aprenen*. Mentre que en altres aplicacions, com en el cas de les xarxes optimitzadores, els pesos i llindars es calculen analíticament en funció del problema a resoldre. A [Bofill et al., 89] es descriuen l'optimització i l'aprenentatge neuronals com dos processos de cerca que tenen lloc a curt i llarg termini, respectivament.

I a [Bofill et al., 90] es presenta una realització del model de retropropagació del gradient, l'algorisme d'aprenentatge neuronal més popular, sobre sistemes multiprocessador.

Aquest treball es centra únicament en les xarxes neuronals optimitzadores (xarxes per resoldre problemes d'optimització) i en aquest apartat es descriuen els models bàsics que fem servir:

- El model discret de Hopfield.
- La màquina de Boltzmann i la recuita simulada.
- Xarxes d'ordre superior.

En aquest treball les xarxes neuronals optimitzadores s'estudien com a tècniques de propòsit general. En conseqüència, no es fa cap intent d'incorporar el coneixement del problema (les propietats dels dissenys) als algorismes de cerca.

### 1.2.1 El model discret de Hopfield i la cerca descendent

El model discret de Hopfield [Hopfield, 82] realitza una estratègia de cerca local descendent sobre un espai de variables binàries. Les unitats de procés de la xarxa són les variables d'estat ( $x_i \in \{0, 1\}, i = 1, \dots, d$ ) i el seu valor determina un *estat* o *configuració* de la xarxa en un instant determinat. L'espai de cerca ( $2^d$  configuracions possibles) són els vèrtexos d'un hipercub binari. L'estratègia de relaxació consisteix en partir d'un estat inicial a l'atzar i actualitzar cada unitat (de manera seqüencial o asíncrona) segons la *regla de decisió* següent:

$$x_i \leftarrow \begin{cases} 1, & \text{si } \phi_i < 0 \\ 0, & \text{altrament,} \end{cases} \quad (1.10)$$

amb el camp local  $\phi_i$  donat per (1.8). Si el nou valor  $x_i^n$  de la unitat  $i$  és diferent del vell direm que s'ha produït una *transició* ( $x_i \rightarrow \bar{x}_i$ ), i el sistema realitza un moviment en la direcció  $i$  de l'espai d'estats. L'algorisme *convergeix* quan no es pot moure en cap de les  $d$  direccions i el valor de totes les variables és *consistent* amb el que els assigna la regla de decisió. La funció  $f$  de l'equació (1.9), en aquest cas, és la funció graó i, sense pèrdua de generalitat, la definim en sentit descendent per coherència amb el conveni de signes que s'utilitza a la resta del treball.

Si les connexions són simètriques ( $w_{ij} = w_{ji}$ ), la *convergència* de l'algorisme està garantida perquè és possible definir una funció d'*energia computacional*

$$E = \frac{1}{2} \sum_{(i,j)} w_{ij} x_i x_j - \sum_i \theta_i x_i, \quad (1.11)$$

que associa a cada estat la suma dels pesos de les connexions i l'indica actius (que connecten unitats en estat 1). L'*increment* d'energia respecte a la unitat  $i$  és, aleshores,

$$\Delta^i E = E|_{x_i=1} - E|_{x_i=0} = \phi_i, \quad (1.12)$$

i, amb  $x_i$  el valor actual, l'energia de transició ve definida per

$$\Delta_{trans}^i E = E|_{\bar{x}_i} - E|_{x_i} = (1 - 2x_i)\phi_i. \quad (1.13)$$

En termes de transicions, la regla de decisió anterior es reescriu aleshores com

$$x_i \leftarrow \bar{x}_i, \text{ si } \Delta_{trans}^i E < 0. \quad (1.14)$$

Amb aquesta regla, doncs, el sistema només accepta les transicions que fan disminuir l'energia (cerca *descendent*). I, com que l'energia és acotada, l'algorisme convergeix necessàriament en un mínim.

Donat que totes les transicions es realitzen entre configuracions veïnes (entre configuracions que només es diferencien en l'estat d'una sola unitat), diem que el model de Hopfield realitza cerca *local*. L'avantatge principal que s'en deriva és que el nou valor  $E^n$  de l'energia, en cas de produir-se una transició, es pot avaluar *incrementalment* com

$$E^n = E + \Delta_{trans}^i E. \quad (1.15)$$

I el cost computacional de l'avaluació de  $\Delta E_{trans}$  (el còmput del camp local a l'eq. 1.8) és inferior en un ordre al de l'avaluació directa d' $E$  a l'equació (1.11). En termes de relaxació, la convergència del sistema es produeix sempre en algun mínim *local*. Però, i aquest és l'inconvenient principal del model, quan el *paisatge d'energies* no és convex, el mínim obtingut no és necessàriament *global*. No ens garanteix, per tant, que haguem trobat el punt *òptim* que minimitza  $E$ .

El model de Hopfield va ser proposat inicialment per a aplicacions de memòria associativa (els mínims de l'energia són els estats que es volen evocar), i a [Hopfield, 84] el model es va estendre a valors dels  $x_i$  continus. A [Hopfield & Tank, 85] es presenta la seva aplicació a problemes d'optimització combinatòria, tema que es descriu en profunditat a [Aarts & Korst, 87]. La idea bàsica és associar les unitats de procés de la xarxa a les variables d'estat del problema a resoldre, i escollir els pesos de les connexions i llindars de tal manera que els mínims d'energia de la xarxa es corresponguin biunívocament amb els òptims de la funció de cost del problema. A [Jagota & Regan, 97] es descriu una aplicació recent d'aquesta tècnica.

Com veurem en el proper apartat, l'ús del concepte d'energia computacional introdueix un paral·lelisme entre una xarxa optimitzadora i alguns models de la mecànica estadística.

### 1.2.2 La màquina de Boltzmann i la recuita simulada

La recuita simulada [Kirkpatrick et al., 83] és una estratègia per introduir aleatorietat al procés de cerca descendent amb l'objectiu d'evitar mínims locals i, combinada amb el model de Hopfield, es coneix com a màquina de Boltzmann [Ackley et al., 85]. En aquest model, enlloc d'actualitzar de manera determinista cada variable d'estat, el que es fa és assignar-li un valor amb una determinada probabilitat, segons una nova

regla de decisió *estocàstica*. En termes de transicions, la probabilitat  $P\{x_i \rightarrow \bar{x}_i\}$  de commutar la unitat  $i$  s'assigna pel mètode de Montecarlo d'acord amb

$$P\{x_i \rightarrow \bar{x}_i\} = g(\Delta^i E_{trans}; T), \quad (1.16)$$

on la funció d'*acotació*  $g : \mathcal{R} \rightarrow [0, 1]$  que s'utilitza amb més freqüència és o bé l'anomenada funció *sigmoïdal*

$$g(z; T) = \frac{1}{1 + e^{\frac{z}{T}}}, \quad (1.17)$$

o bé la l'anomenada *regla de Metropolis* [Metropolis et al., 53], que en la nostra formulació s'expressa com

$$g(z; T) = \begin{cases} 1 & \text{si } z < 0 \\ e^{-\frac{z}{T}} & \text{altrament.} \end{cases} \quad (1.18)$$

En aquest treball hem optat per la segona, perquè condueix a una relaxació més ràpida. La regla de Metropolis accepta totes les transicions que fan decreïxer l'energia de la xarxa, però també accepta transicions ascendents amb una probabilitat que depèn de la *temperatura computacional*  $T$ . Aquest paràmetre s'utilitza aleshores per controlar el nivell d'aleatorietat. Per a  $T \rightarrow \infty$  s'accepten *totes* les transicions, i per a  $T \rightarrow 0$  el model es redueix al de Hopfield.

La màquina de Boltzmann és un model estocàstic que es pot descriure en termes de mecànica estadística (vegeu, p.ex., [Hertz et al., 93],) i la seva dinàmica ve governada per l'*energia lliure* del sistema  $F$ ,

$$F = \langle E \rangle - TS, \quad (1.19)$$

on  $\langle E \rangle$  és l'energia mitjana i  $S$  és l'*entropia*. Quan s'assoleix l'*equilibri tèrmic* l'energia lliure  $F$  es minimitza, i totes les variables d'estat es fan *estacionàries*. En particular, la probabilitat  $P_\alpha$  de trobar el sistema en un estat  $\alpha$  qualsevol d'energia  $E_\alpha$  es fa independent del temps, i segueix una distribució de Boltzmann-Gibbs (d'aquí el nom del model,)

$$P_\alpha = \frac{1}{Z} e^{-\frac{E_\alpha}{T}}, \quad (1.20)$$

on  $Z$  és una constant de normalització (la *funció de partició*.) Energia mitjana i entropia es poden expressar aleshores com

$$\langle E \rangle = \sum_{\alpha} E_{\alpha} P_{\alpha}, \quad (1.21)$$

$$S = - \sum_{\alpha} P_{\alpha} \log P_{\alpha}, \quad (1.22)$$

i

$$F = \sum_{\alpha} E_{\alpha} P_{\alpha} + T \sum_{\alpha} P_{\alpha} \log P_{\alpha}, \quad (1.23)$$

és l'energia lliure de la distribució.

Una condició suficient perquè un sistema dinàmic es mantingui en equilibri tèrmic és que, en mitjana, el nombre de transicions entre dos estats  $\alpha$  i  $\beta$  qualsevols sigui el mateix. Es a dir, que

$$P_\alpha P\{\alpha \rightarrow \beta\} = P_\beta P\{\beta \rightarrow \alpha\}, \quad (1.24)$$

on  $P\{\alpha \rightarrow \beta\}$  és la probabilitat de que es produeixi una transició de l'estat  $\alpha$  a l'estat  $\beta$ . En el cas de la distribució de Boltzmann-Gibbs, si reescribim l'equació anterior en forma de cocients, i tenint en compte (1.20), la condició d'equilibri es pot reescriure com

$$\frac{P\{\alpha \rightarrow \beta\}}{P\{\beta \rightarrow \alpha\}} = \frac{P_\beta}{P_\alpha} = e^{-\frac{E_\beta - E_\alpha}{T}}. \quad (1.25)$$

Com que a la màquina de Boltzmann només tenen probabilitat diferent de zero les transicions entre estats veïns, si dos d'aquests estats es diferencien en la unitat  $i$  i associem  $\alpha$  al valor  $x_i$  i  $\beta$  al valor  $\bar{x}_i$ , aleshores  $P\{\alpha \rightarrow \beta\} = P\{x_i \rightarrow \bar{x}_i\}$  que, desenvolupant per a qualsevol de les dos expressions de  $g$  a (1.16), condueix a

$$\frac{P\{x_i \rightarrow \bar{x}_i\}}{P\{\bar{x}_i \rightarrow x_i\}} = e^{-\frac{\Delta_{\text{trans}}^i E}{T}}. \quad (1.26)$$

La regla de decisió garanteix, doncs, que la situació d'equilibri tèrmic a la màquina de Boltzmann és estable.

La distribució de probabilitats de l'equació (1.25) és important perquè indica que l'aleatorietat introduïda no és arbitrària. Per a dos estats  $\alpha$  i  $\beta$  qualsevols, la relació de probabilitats depèn de la diferència d'energies, ponderada per la inversa de  $T$ . D'aquesta manera, els estats de menor energia són més probables i el sistema té tendència a evolucionar per les zones en que l'energia mitjana és més baixa. El paràmetre  $T$ , aleshores, permet controlar la profunditat de les "valls" i "barreres" del paisatge tèrmic resultant.

L'estratègia de la *recuita simulada* consisteix en començar a una temperatura prou elevada i anar "refredant" poc a poc. Un refredament massa ràpid congelaria el sistema en un mínim local. Però si el refredament és prou lent, i les perturbacions a l'equilibri tèrmic no són molt importants, el sistema evoluciona cap a valls cada cop més profundes, i la probabilitat de trobar un mínim global augmenta. Es demostra que, en el límit (refredament infinitament lent), l'algorisme convergeix asimptòticament cap a un òptim (vegeu, p.e, [Aarts & Korst, 88]). A la pràctica, amb temps finit, el model és molt eficaç per trobar *pseudo*-òptims de qualitat.

La dificultat principal de la recuita simulada està en trobar una llei de decrement de la temperatura adequada al problema. El més habitual en la literatura és utilitzar una llei geomètrica del tipus

$$T_n = \tau T_{n-1}, \quad (1.27)$$

on  $\tau < 1$  és una constant que regula la velocitat de decaïment (típicament  $\tau = 0.8 \dots 0.99$ ). Les temperatures inicial  $T_0$  i final  $T_f$  també s'han d'escollir amb compte

per no perdre el temps iterant a temperatures massa altes (on tots els estats són pràcticament equiprobables) ni massa baixes (quan el sistema ja ha quedat “atrapat” en una vall). El que convé, doncs, és invertir tot el temps disponible en refredar lentament al voltant de la *temperatura crítica*  $T_{cri}$  a la qual es formen les valls significatives (vegeu, p.e., [Hertz et al., 93]).

Tot i ser un model teòricament atractiu, l’inconvenient de la màquina de Boltzmann és que el seu cost computacional és elevat. L’extensió natural del model que acabem de descriure és la màquina de Boltzmann determinista [Peterson & Anderson, 87, Aarts & Korst, 88, Hertz et al., 93] que, basada en la teoria del camp mitjà de la mecànica estadística, enlloc d’explorar la distribució d’equilibri pel mètode de Montecarlo, calcula les mitjanes tèrmiques corresponents a cada temperatura. La seva realització requereix variables reals i per aquest motiu, com veurem més endavant, no ha estat inclosa en el nostre repertori d’algorismes de cerca. Algunes variants i aplicacions d’aquest model es troben, per exemple, a [Peterson & Södeberg, 89, Gutzmann, 87].

La màquina de Boltzmann s’utilitza també en aplicacions d’associació de patrons en les que, a partir d’un conjunt d’entrenament, els pesos s’aprenen segons un criteri de màxima versamblança entre les distribució forçada i lliure [Hinton et al., 84].

### 1.2.3 Xarxes d’ordre superior

Els models de xarxa neuronal que hem considerat fins ara només tenen en compte les restriccions d’aritat 1 (els llindars) i 2 (les connexions). El model és directament extensible a restriccions de qualsevol aritat (xarxes d’*ordre superior*) [Sejnowski, 86]. Les equacions (1.8) i (1.9) es reescriuen aleshores com

$$\phi_i = -\theta_i + \sum_j w_{ij}x_j + \sum_{(j,k)} w_{ijk}x_jx_k + \sum_{(j,k,l)} w_{ijkl}x_jx_kx_l \cdots, \quad (1.28)$$

$$x_i \leftarrow f(\phi_i), \quad (1.29)$$

on  $w_{ijk}$ ,  $w_{ijkl}$ , ..., són els pesos de les connexions d’aritat 3, 4, etc. (No s’admeten subíndexos repetits).

Per a aplicacions d’optimització (vegeu, p.ex., [Joya et al., 91]), definim  $E$  com la suma dels pesos de les totes les connexions actives

$$E = -\sum_i \theta_i x_i + \sum_{(i,j)} w_{ij}x_i x_j + \sum_{(i,j,k)} w_{ijk}x_i x_j x_k + \sum_{(i,j,k,l)} w_{ijkl}x_i x_j x_k x_l + \cdots, \quad (1.30)$$

on una connexió està activa quan totes les variables que hi intervenen ho estan. Si assumim que les connexions són simètriques (el valor dels pesos no depèn de l’ordre dels subíndexos), aleshores  $E$  és una funció d’energia com en el model de Hopfield, i les regles de decisió descendent (eq 1.14) i estocàstica (eq. 1.16) es poden aplicar directament. El camp local a la unitat  $i$  correspon aleshores, com abans, a l’increment d’energia  $\Delta^i E$ . En aquest treball, els models de xarxa que farem servir tenen restriccions d’aritat 1, 2 i 4.



La dinàmica de les xarxes optimitzadores s'estudia a [Goles & Matamala, 94] i a [Jagota & Wang, 96], entre d'altres. A [Van den Berg, 96] es proposa una extensió del model de Hopfield amb una definició de l'energia lliure que admet expressions generalitzades tan per als camps locals (xarxes d'ordre superior incluídes), com per a la funció d'acotació (eq. 1.16).

### 1.3 Objectius de la tesi i organització de la memòria

L'objectiu principal d'aquest treball és l'*avaluació* de les xarxes neuronals optimitzadores, enteses com *algorisme* de cerca, enfront d'un *problema* difícil, la generació de dissenys en blocs. Cada instància del problema és un  $DB(v, b, u)$ .

El primer pas consisteix en la formulació del problema en termes d'optimització combinatòria, és a dir, en definir *funcions de cost*  $F$ , els mínims globals de les quals corresponguin a solucions del problema. Les funcions que es proposen en aquest treball es basen en el *nombre de parelles* com a mesura de distribució de cada una de les propietats d'un disseny: nombre d'uns, uns per fila, uns per columna i correlacions entre parelles de files que, en el darrer cas, condueixen a restriccions de quart ordre. Les funcions que s'obtenen, aleshores, venen determinades per la seva *estructura*, que defineix quines d'aquestes propietats intervenen, i els coeficients que determinen la seva *composició* relativa. El domini d'aplicació de la funció és l'espai de configuracions- $v \times b$  associat a cada instància del problema, i els coeficients de composició són paramètrics en  $(v, b, u)$ . En la formulació que fem servir, cada funció es *projecta* de manera natural sobre una xarxa optimitzadora de dimensions  $v \times b$ , amb connexions d'ordre 1, 2 i 4, per a cada tipus de restricció. Cada unitat de la xarxa correspon a una variable d'estat, i els pesos de les connexions corresponen als coeficients de composició, de manera que l'energia  $E$  de cada estat de la xarxa correspon exactament al valor de la funció. *Xarxa* i *funció* són, aleshores, completament equivalents. Un segon grup de funcions, que denominarem de tipus  $H$ , es basen en les propietats del disseny per fer minimització de segon ordre (minimització quadràtica entre els camps locals actual i òptim). La seva estructura interna no s'adiu tan bé a les xarxes, però es poden formular equivalències.

Els models neuronals descrits a l'apartat anterior comparteixen l'estructura de la xarxa (unitats i connexions), i únicament es diferencien en l'*estratègia* de cerca que fan servir. *Cerca descendent* CD, en el cas del model de Hopfield, i *recuita simulada* RS, en el cas de la màquina de Boltzmann. RS és paramètrica, en termes de les temperatures inicial i final, i el ritme de refredament. A continuació es defineix una formulació algorísmica de la cerca local, que descomposa l'algorisme de cerca en un subconjunt d'operacions relacionades amb la xarxa (o la funció) i un subconjunt d'operacions relacionades amb l'estratègia. En aquest marc es reformulen les dues estratègies anteriors, i es proposa una estratègia nova, la *cerca cooperativa* CC, basada en l'exploració paral·lela de l'espai de cerca per mitjà de diverses instàncies

de xarxa evolucionant *en esquadra* cap a zones d'energia mitjana decreixent. La reducció progressiva del radi de l'esquadra condueix el sistema a valls cada vegada més profundes. Es proposen tres variants de l'estratègia que, com en el cas d'RS, és paramètrica (tamany  $S$  de l'esquadra, radis inicial i final, i ritme de decrement).

L'algorisme anterior es simula en un entorn multiprocessador SIMD basat en una Connection Machine CM2 [TMC], un hipercub amb  $N = 2048$  processadors d'un bit, que s'adiu perfectament a la naturalesa binària de les unitats neuronals. Aquest és el motiu pel qual, com s'ha dit, no hem inclòs la màquina de Boltzmann determinista en el nostre repertori. Però es tracta també d'una de les motivacions principals en la definició de la nova estratègia. Amb un cost de comunicació molt baix, la cerca cooperativa és un algorisme paral·lel que pretén treure el màxim rendiment de diversos processadors treballant de manera *cooperativa* per a la resolució d'un mateix problema.

L'avaluació d'aquests algorismes es farà aleshores combinant les diferents funcions de cost amb les tres estratègies de cerca, aplicant-els a un repertori de dissenys de prova que no siguin massa grans. El problema resultant (el problema de l'avaluació), és un cas no trivial d'*anàlisi experimental*, i s'ha d'abordar amb precaució [Montgomery, 91]. Problema, funció i estratègia s'identifiquen com els *factors experimentals* principals. Mentre que l'estructura i composició de les funcions, o les variants i paràmetres de les estratègies corresponen a subfactors. L'anàlisi factorial complet és aleshores intractable. En una primera etapa d'*entrenament*, amb un subconjunt de problemes, s'ajusten composicions, es seleccionen estructures, i es determinen els paràmetres i variants, en un procés d'optimització experimental. En una segona etapa es fan les *comparacions* per a seleccionar les millors funcions i les estratègies més eficients, i s'analitza l'evolució dels resultats amb el tamany del problema. En una tercera etapa, s'aplica el millor algorisme trobat a problemes cada vegada més grans.

La variable escollida, com a *resultat experimental* és el *cost de relaxació z fins la primera solució* que medeix l'*eficiència* de la cerca, ja que té en compte, de manera integrada, l'*eficàcia* de l'algorisme per trobar solucions, i els *recursos* invertits en aconseguir-ho. En l'entorn d'execució presentat, però, un *experiment paral·lel* és l'execució simultània d'un descens a cada processador.  $E(z)$  s'estima aleshores per mitjà de la variable  $w$  (cost total / nombre de solucions trobades), que es defineix com *resposta* de l'experiment paral·lel. Donat que la generació de dissenys és NP-completa, un indicador important per a l'avaluació d'algorismes és el tamany màxim dels problemes que l'algorisme és capaç de resoldre.

La formulació utilitzada en la definició de les funcions de cost posa de manifest algunes propietats noves dels dissenys en blocs, en termes de les mesures que es fan servir, i porta de manera natural a l'extensió del concepte de disseny a un nou tipus de configuracions combinatòries que anomenarem *dissenyos màximament equilibrats*, DMEs. Donat un conjunt  $(v, b, u)$  de paràmetres, els DMEs accepten fins a dos valors diferents en el nombre d'uns per fila ( $r$  i  $r + 1$ ), dos valors diferents en el nombre d'uns per columna ( $k$  i  $k + 1$ ), i dos valors diferents en les correlacions entre

files ( $\lambda$  i  $\lambda + 1$ ). Els DMEs constitueixen, aleshores, un super-conjunt dels DBs, per als quals les restriccions d'admissibilitat s'han suprimit. A continuació, s'analitzen algunes de les seves propietats, i s'aplica la millor estratègia de la fase experimental a la seva generació.

Recapitulant, els objectius que es formulen en aquest treball són els següents:

- Avaluació de xarxes neuronals optimitadores per a la generació de dissenys en blocs:
  - Proposta i formulació de funcions de cost per als DBs.
  - Projecció de les funcions de cost sobre una xarxa optimitzadora.
  - Proposta i formulació d'estratègies de cerca.
  - Avaluació experimental:
    - \* Entrenament i comparació entre problemes, funcions i estratègies, per tal de determinar el millor algorisme sobre un conjunt de problemes de prova.
    - \* Comportament del millor algorisme trobat davant de problemes cada vegada més grans.
  - Extensió al cas dels dissenys màximament equilibrats.
- Anàlisi d'algunes propietats de DBs i DMEs, en termes de la formulació utilitzada.

L'organització de la memòria és la següent. El capítol 2 correspon a la definició de les funcions de cost i la seva projecció a la xarxa. Es defineix el nombre de parelles com a mesura de distribució, s'aplica a les propietats d'un disseny i s'analitza la condició d'optimalitat (valor de les mesures en una configuració òptima i requisits mínims d'optimalitat). Les mesures obtingudes es combinen, s'imposen condicions de validesa, i s'escull un repertori de funcions. La projecció sobre una xarxa d'ordre 4 és immediata. En el darrer apartat, s'explota el coneixement adquirit sobre les propietats d'un disseny per a definir les funcions de tipus  $H$ , i s'analitza el camp local equivalent.

El capítol 3 presenta les estratègies. Es defineix el marc global i es formulen les dues estratègies clàssiques. A continuació es presenta i formalitza la cerca cooperativa, juntament amb una petita anàlisi de les seves propietats.

El capítol 4 formalitza el procés d'avaluació experimental. Per a un cas experimental particular (un problema i un algorisme), es defineix el *cost de relaxació* (en iteracions), per contraposició al temps d'execució, depenent de la realització, i a la complexitat computacional (de cada iteració). I es caracteritzen les variables resultat  $z$  i resposta  $w$  en termes del resultat  $x = \{\text{èxit}, \text{fracàs}\}$  i el cost  $c$  de l'experiment *elemental* (un descens). Finalment, es fa una valoració de l'espai factorial d'experimentació i es planifiquen els experiments.

El capítol 5 correspon a la realització dels experiments. En l'etapa d'entrenament, es caracteritzen estadísticament les variables resultat  $z$  i resposta  $w$ , es determinen les millors funcions (estructura i composició), i s'ajusten els paràmetres de les estratègies. En la segona etapa es fan les comparacions, i s'estudia l'evolució dels resultats en funció del tamany del problema, i en la darrera etapa s'aplica el millor algorisme obtingut a problemes cada vegada més grans.

En el capítol 6 es formulen els dissenys màximament equilibrats, s'estudien algunes de les seves propietats, i es revisa quines funcions de cost del repertori són vàlides per a un DME qualsevol (pitjor cas). La casuística de DMEs s'avalua també experimentalment i, en un experiment final, s'aplica l'algorisme de cerca sobre un conjunt de problemes de prova.

Per acabar, en el capítol 7, es revisen les aportacions del treball i es sistematitzen les conclusions de l'avaluació experimental. La discussió de les tècniques emprades condueix a la formulació de perspectives de futur.

## Capítol 2

# Funcions de cost per a la generació de dissenys amb xarxes optimitzadores

L'objectiu d'aquest capítol és l'elecció de funcions de cost *vàlides* per a la generació de dissenys en blocs i *estructuralment adequades* per a la seva realització amb xarxes optimitzadores.

Fixada la notació, es defineixen *mesures de distribució* per a cada una de les propietats d'un disseny (nombre d'uns, uns per fila, uns per columna i correlacions entre files), i del disseny complementari. La mesura bàsica que s'utilitza és el *nombre de parelles*, que s'il·lustra prèviament amb l'anàlisi de la distribució de  $w$  boles en  $p$  pots de capacitat  $q$ . Aplicada al cas de les correlacions, aquesta mesura condueix a restriccions de quart ordre que anomenarem *quàdruples*. Definim el valor *òptim* d'una mesura com la seva cota inferior en una configuració hipotètica que compleixi les propietats d'un disseny. Nombre de quàdruples òptim i nombre d'uns just són, aleshores, condició necessària i suficient d'optimalitat i, conjuntament, constitueixen la definició més breu de disseny en termes de les variables d'estat.

Definim una funció de cost  $F$  per a la generació de dissenys en blocs com una combinació lineal de les mesures anteriors que satisfaci les *condicions de validesa* (en termes dels seus coeficients). L'*estructura* d'una funció es defineix aleshores pels termes que hi apareixen amb coeficient no nul, i la seva *composició* pel valor de la resta de coeficients. El repertori de funcions resultant es projecta, finalment, sobre una xarxa optimitzadora definida amb la mateixa estructura: una unitat per a cada variable d'estat, i un tipus de connexió per a cada mesura. Així, si s'ajusten els pesos al valor dels coeficients, l'energia de la xarxa i el cost de la funció s'identifiquen.

En el darrer apartat, coneixent les propietats d'un disseny, es defineix l'error quadràtic entre els camps locals òptim i actual com a objectiu alternatiu de minimització. Això condueix a un nou repertori de funcions, que anomenarem de tipus  $H$ .

## 2.1 Notació i terminologia

Donats  $(v, b, u)$  com a paràmetres independents, una *configuració*  $A - v \times b$  és un arranjament concret dels  $v$  elements en  $b$  blocs, i la describim per la seva matriu d'incidència  $A \equiv [x_{ij}]$ . En termes de cerca, quan l'objectiu és un disseny  $\text{DB}(v, b, u)$ , les entrades  $x_{ij} \in \{0, 1\}$  d'una configuració genèrica són les *variables d'estat*, i l'*espai de cerca* és el conjunt  $\mathcal{A} = \{A\}$  de totes les configuracions- $v \times b$  possibles, de cardinal  $2^{vb}$ . Cas d'existir, qualsevol configuració que satisfaci les propietats del disseny és un *òptim*, amb *descriptors*  $[u, r, k, \lambda]$  fixos.

En sentit ampli, anomenarem *mesura*  $M(A) = M(x_{11}, x_{12}, \dots, x_{vb})$  a qualsevol aplicació d' $\mathcal{A}$  en  $\mathcal{Z}$  o  $\mathcal{R}$  que assigni un valor a cada configuració. Una *mesura*  $M$  és *vàlida* per a medir una determinada propietat si és capaç de discriminar el grau en que diferents configuracions satisfan aquesta propietat.

Direm que dues configuracions  $A$  i  $A'$  són *veïnes* si només es diferencien en el valor d'un únic component  $(i, j)$ . Sigui  $x_{ij}$  el seu valor en  $A$ , i sigui  $\bar{x}_{ij} = 1 - x_{ij}$  el seu valor en  $A'$ . Definim l'*increment local* d' $M$  respecte a  $x_{ij}$  com

$$\Delta^{ij}M(A) = M|_{x_{ij}=1} - M|_{x_{ij}=0} = \begin{cases} M(A) - M(A'), & \text{quan } x_{ij} = 1 \\ M(A') - M(A), & \text{quan } x_{ij} = 0, \end{cases} \quad (2.1)$$

i definim l'*increment de transició* d' $M$  respecte a  $x_{ij}$  com

$$\Delta_{trans}^{ij}M(A) = M|_{\bar{x}_{ij}} - M|_{x_{ij}} = M(A') - M(A). \quad (2.2)$$

La relació entre els dos increments és aleshores

$$\Delta_{trans}^{ij}M(A) = (1 - 2x_{ij})\Delta^{ij}M(A). \quad (2.3)$$

La *mesura més elemental* és  $U$ , que compta el nombre d'uns d'una configuració

$$U(A) = \sum_{i=1}^v \sum_{j=1}^b x_{ij}. \quad (2.4)$$

Sigui  $U(A) = o^1$ , i sigui  $\mathcal{A}_o \subset \mathcal{A}$  el subconjunt de configuracions que tenen exactament  $o$  uns,  $\mathcal{A}_o \equiv \{A : U(A) = o\}$ . Per a una *mesura*  $M$  qualsevol definim

$$(M|U)^{\min}(o) = \min_{\mathcal{A}_o} M(A) \quad (2.5)$$

$$(M|U)^{\max}(o) = \max_{\mathcal{A}_o} M(A), \quad (2.6)$$

com les cotes inferior i superior d' $M$  *condicionades al nombre d'uns*. I definim l'*increment* d'aquestes cotes com,

$$\Delta(M|U)^{\min}(o) = (M|U)^{\min}(o+1) - (M|U)^{\min}(o) \quad (2.7)$$

$$\Delta(M|U)^{\max}(o) = (M|U)^{\max}(o+1) - (M|U)^{\max}(o). \quad (2.8)$$

---

<sup>1</sup>El símbol  $u$ , que és un paràmetre, denota el nombre d'uns en el *disseny*

El mínim  $M^{\min}$  de la cota inferior és el mínim absolut de la mesura, i el màxim  $M^{\max}$  de la cota superior és el màxim absolut. Direm que una mesura té un mínim (o màxim) *puntual*, si aquest mínim (o màxim) absolut es produeix per a un únic valor d' $o$ , que denotarem  $o^*$ .

Per a qualsevol mesura  $n = N(A)$  que estableixi una partició en  $\mathcal{A}$ , les cotes inferior i superior d' $M$  condicionades a  $N$  es defineixen de la mateixa manera, i les denotarem per  $(M|N)^{\min}(n)$  i  $(M|N)^{\max}(n)$ , respectivament.

Finalment, sobre el mateix espai  $\mathcal{A}$ , els símbols sobre-ratllats denoten referències a la configuració  $\bar{A} \equiv [\bar{x}_{ij}]$ , *complementària* d' $A$ . Així, si  $o$  denota el nombre d'uns a  $A$ ,  $\bar{o}$  denota el nombre d'uns a  $\bar{A}$ , amb  $\bar{o} = vb - o$ . I si  $(v, b, u)$  són els paràmetres d'un disseny, aleshores  $(v, b, \bar{u})$  són els paràmetres del disseny complementari, amb descriptors  $(\bar{u}, \bar{r}, \bar{k}, \bar{\lambda})$ .

Donada una mesura  $M$  sobre  $\mathcal{A}$ , definim la *mesura complementària*  $\bar{M}$  com

$$\bar{M}(A) = M(\bar{A}). \quad (2.9)$$

Les seves cotes i increments compleixen, aleshores, les relacions següents:

$$(\bar{M}|U)^{\min}(o) = (M|U)^{\min}(\bar{o}) \quad (2.10)$$

$$(\bar{M}|U)^{\max}(o) = (M|U)^{\max}(\bar{o}), \quad (2.11)$$

i

$$\Delta(\bar{M}|U)^{\min}(o) = -\Delta(M|U)^{\min}(\bar{o} - 1) \quad (2.12)$$

$$\Delta(\bar{M}|U)^{\max}(o) = -\Delta(M|U)^{\max}(\bar{o} - 1). \quad (2.13)$$

## 2.2 Mesures de distribució per a les propietats d'un disseny

Com s'ha vist a l'apartat 1.1, les propietats que caracteritzen la matriu d'incidència d'un disseny  $DB(v, b, u)$  són el seu nombre d'uns, la uniformitat de files i columnes, i l'equilibri en les correlacions entre parelles de files. En aquest apartat es defineix el nombre de parelles aplicat, de manera genèrica, a la mesura de la distribució de  $w$  boles en  $p$  pots de capacitat  $q$ , i s'extén, a continuació, a les propietats d'un disseny.

### 2.2.1 Mesures de distribució basades en el nombre de parelles

L'arranjament d'un nombre indeterminat de boles en  $p$  pots depèn del nombre total de boles  $w$ , i de com estan repartides entre els pots. Amb independència del nombre concret de boles, direm que la *distribució* de les boles en els pots és *uniforme* quan el nombre de boles a cada pot és el més similar possible. I direm que el *pitjor cas* en la distribució es produeix quan totes les boles ocupen el mínim nombre de pots possible.

**Definició 2.1** *Sigui  $p$  el nombre de pots i sigui  $b_j$  el nombre de boles al pot  $j$  ( $j = 1, \dots, p$ ). Direm que les boles estan màximament o uniformement distribuïdes entre els pots quan, per a algun valor fix  $b$ ,*

$$(b_j = b) \text{ o } (b_j = b + 1), \quad \forall j. \quad (2.14)$$

En aquestes circumstàncies,  $w = \sum_{j=1}^p b_j$  és el nombre total de boles i, amb  $b = \lfloor w/p \rfloor$  i  $p_0 = w \bmod p$ , tindrem  $p_0$  pots amb  $b + 1$  boles, i la resta amb  $b$ . Els valors  $b$  i  $p_0$ , que descriuen la distribució uniforme, depenen doncs del nombre total de boles, i els expressarem com  $b(w)$  i  $p_0(w)$ . Per a  $(p, w)$  tals que  $w|p$  (condició de multiplicitat) es compleix que  $p_0(w) = 0$  i cada pot té exactament  $b(w)$  boles. Direm aleshores que la distribució és *estrictament* uniforme.

**Definició 2.2** *Sigui  $p$  el nombre de pots i  $q$  la capacitat de cada pot. Sigui  $b_j$  el nombre de boles al pot  $j$  ( $j = 1, \dots, p$ ). Direm que les boles segueixen una distribució de pitjor cas quan, com a màxim, existeix un  $j$  per al qual*

$$b_j \notin \{0, q\}. \quad (2.15)$$

En aquestes circumstàncies, si  $w$  és el nombre de boles com abans,  $p_1(w) = \lfloor w/q \rfloor$  i  $b_1(w) = w \bmod q$ , aleshores tindrem  $p_1(w)$  pots completament plens i un pot amb  $b_1(w)$  boles. La resta de pots estaran buits.

El nombre total de boles  $W(B)$  estableix una partició en l'espai  $\mathcal{B} = \{B\}$  de totes les configuracions possibles amb  $p$  pots de capacitat  $q$ , en funció del seu nombre de boles  $w$ . Una mesura  $D : \mathcal{B} \rightarrow \mathcal{R}$  medeix la *distribució* de les boles si les seves cotes condicionades a  $W$  corresponen als dos casos extrems de la distribució. Formalment,

**Definició 2.3** *Direm que  $D$  és una mesura vàlida de la distribució de  $0 \leq w \leq qp$  boles en  $p$  pots de capacitat  $q$  si i només si, per a qualsevol configuració  $B$  tal que  $W(B) = w$ , es compleix que*

$$D(B) = (D|W)^{\min}(w) \Leftrightarrow B \text{ té distribució uniforme, i} \quad (2.16)$$

$$D(B) = (D|W)^{\max}(w) \Leftrightarrow B \text{ té distribució de pitjor cas.} \quad (2.17)$$

**Proposició 2.1** *El nombre total de parelles que es poden formar agrupant boles pertanyents a un mateix pot,*

$$P(B) = \sum_j \binom{b_j}{2}, \quad (2.18)$$

*és una mesura vàlida de la distribució de  $0 \leq w \leq qp$  boles en  $p$  pots de capacitat  $q$ .*

Fixat el nombre de boles, es verifica fàcilment que el nombre de parelles és mínim o màxim quan la distribució és uniforme o de pitjor cas, respectivament. De manera



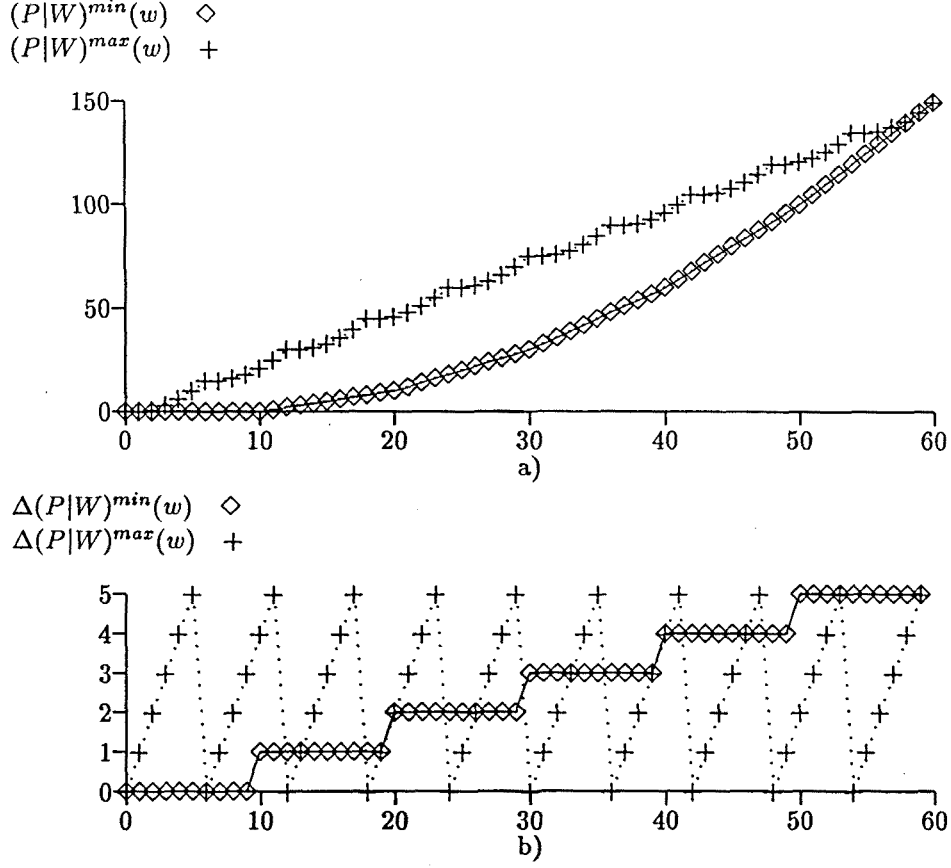


Figura 2.1: a) Cotes inferior i superior de  $P(B)$  en funció del nombre de boles i b) increments corresponents, per al cas  $q = 6$ ,  $p = 10$ .

que, a partir de les definicions 2.1 i 2.2, les cotes inferior i superior de  $P(B)$  són

$$(P|W)^{\min}(w) = p \binom{b(w)}{2} + p_0(w)b(w), \quad (2.19)$$

$$(P|W)^{\max}(w) = p_1(w) \binom{q}{2} + \binom{b_1(w)}{2}. \quad (2.20)$$

Els seus increments respectius valen

$$\Delta(P|W)^{\min}(w) = b(w) \quad (2.21)$$

$$\Delta(P|W)^{\max}(w) = b_1(w). \quad (2.22)$$

La figura 2.1 mostra les cotes del nombre de parelles, així com els seus increments, per al cas en que  $q = 6$  i  $p = 10$ .

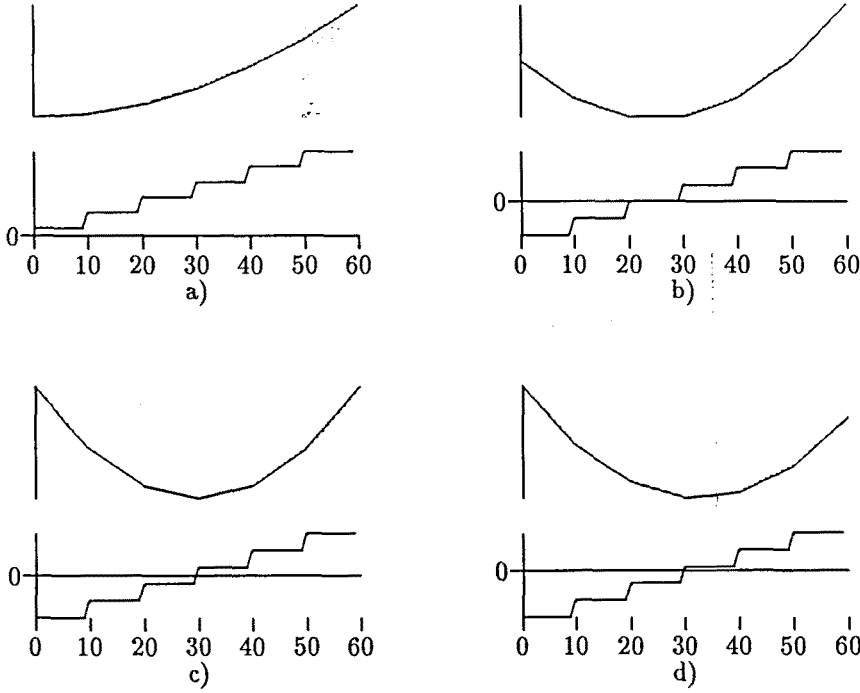


Figura 2.2: Cota inferior i increment de la mesura  $F = \mu_w W + \mu_p P$  amb  $\mu_p = 4$  i a)  $\mu_w = 2$ , b)  $\mu_w = -8$ , c)  $\mu_w = -10$  i d)  $\mu_w = -11$ , per al cas  $q = 6$ ,  $p = 10$ .

Donat que  $W(B)$ , òbviament, no depèn de la distribució de les boles,

$$(W|W)^{\min}(w) = w \text{ i} \quad (2.23)$$

$$\Delta(W|W)^{\min}(w) = 1. \quad (2.24)$$

Aleshores, la combinació lineal

$$F(B) = \mu_w W(B) + \mu_p P(B), \quad (2.25)$$

amb  $\mu_p > 0$ , és una mesura *generalitzada* de la distribució de les boles.

Donat que es tracta de minimitzar, la cota *inferior* d' $F$  condicionada al nombre de boles és

$$(F|W)^{\min}(w) = \mu_w w + \mu_p \left[ p \binom{b(w)}{2} + p_0(w)b(w) \right], \quad (2.26)$$

amb increment

$$\Delta(F|W)^{\min}(w) = \mu_w + \mu_p b(w). \quad (2.27)$$

La figura 2.2 mostra aquestes expressions per a diversos valors de la relació  $\mu_w/\mu_p$  per al cas  $q = 6$  i  $p = 10$ . Pel que fa a l'increment,  $\Delta(F|W)^{\min}(w)$  creix en  $\mu_p$  a cada múltiple de  $p$ , i és constant en els intervals intermitjos. El terme  $\mu_w$  en fixa

l'ordenada a l'origen. De manera que  $(F|W)^{\min}(w)$  és lineal entre múltiples de  $p$  i convexa globalment. Així, el mínim absolut es produeix quan l'increment canvia de signe, i es pot ajustar per mitjà dels coeficients. Com es pot observar, però, únicament si el punt de tall es produeix en un punt  $w^{\text{wast}}$  que satisfaci la condició de multiplicitat  $(w^{\text{wast}}|p)$ , aquest mínim serà *puntual*, i solsament en aquest cas  $F^*$  serà apta a efectes d'optimització. Formalment,

**Definició 2.4** Donat un conjunt de paràmetres  $(q, p, w^*)$ , direm que una mesura  $D : \mathcal{B} \rightarrow \mathcal{R}$  és una funció de cost per a la distribució uniforme d'un nombre prefixat  $w^*$  de boles en  $p$  pots de capacitat  $q$  si i només si es compleixen les condicions següents:

- i) Existeix una cota  $D^*$  tal que  $D(B) \geq D^*, \forall B \in \mathcal{B}$ .
- ii) Per a qualsevol configuració  $B^*$ , amb  $W(B^*) = w^*$  i distribució uniforme, es compleix que  $D(B^*) = D^*$ .
- iii) Per a qualsevol configuració  $B^*$  tal que  $D(B^*) = D^*$ , es compleix que  $W(B^*) = w^*$  i  $B^*$  té distribució uniforme.

Aleshores,

**Proposició 2.2** Sigui  $w^*$  el nombre de boles desitjat. La mesura  $F$  de l'equació (2.25) és una funció de cost per a la distribució uniforme d' $w^*$  boles en  $p$  pots de capacitat  $q$  si i només si

$$w^* = pd, \quad (2.28)$$

amb  $d$  enter, i

$$\Delta(F|W)^{\min}(w^*) > 0 \quad i \quad (2.29)$$

$$\Delta(F|W)^{\min}(w^* - 1) < 0. \quad (2.30)$$

El seu valor òptim és

$$F^* = (F|W)^{\min}(w^*) = \mu_w w^* + \mu_p p \binom{d}{2}. \quad (2.31)$$

En realitat, com hem vist, la condició (2.28) de multiplicitat és condició necessària per a les altres dues, que fixen localment un mínim al nombre de boles desitjat  $w^*$ . Però, donat que  $\Delta(F|W)^{\min}(w)$  és convexa, si les condicions es compleixen serà també el mínim absolut. Això fixa el valor de l'òptim a  $F^*$ , derivat de les equacions (2.23) i (2.19), i per la proposició 2.1, totes i únicament les configuracions amb  $w^*$  boles uniformement distribuïdes l'assoleixen, completant la definició 2.4.

Donats  $(q, p, w^*)$ , aplicant les condicions (2.27, 2.30) a l'expressió de l'increment (eq. 2.27), s'obté el marge de valors permesos de la relació entre els coeficients:

$$d - 1 < -\frac{\mu_w}{\mu_p} < d. \quad (2.32)$$

La condició addicional

$$\Delta(F|W)^{min}(w^* - 1) = -\Delta(F|W)^{min}(w^*), \quad (2.33)$$

que condueix a

$$\frac{\mu_w}{\mu_p} = -d + \frac{1}{2}, \quad (2.34)$$

correspon al que anomenarem una funció de cost *simètrica*. La figura 2.2c) n'es un exemple.

Dues formulacions específiques de mesures simètriques són la suma de quadrats  $Qua(B)$ , i l'error quadràtic  $Err^d(B)$  respecte a un nombre desitjat  $d$  de boles a cada pot.

$$\begin{aligned} Qua(B) &= \sum_j b_j^2 = \\ &= W(B) + 2P(B) \end{aligned} \quad (2.35)$$

$$\begin{aligned} Err^d(B) &= \sum_j (b_j - d)^2 = \\ &= (1 - 2d)W(B) + 2P(B) + pd^2. \end{aligned} \quad (2.36)$$

Com s'observa, prescindint de la constant  $pd^2$ , que no afecta l'optimització, les dues expressions són casos particulars de la mesura  $F$  en que els coeficients  $(\mu_w, \mu_p)$  valen  $(1, 2)$  i  $(1 - 2d, 2)$ , respectivament. La segona és mínima per a  $w^* = pd$ . La primera, el cas extrem de l'expressió (2.34), és mínima per a  $w^* = d = 0$ .

Finalment, donada una configuració  $B$ , les mesures complementàries  $\bar{W}(B)$  i  $\bar{P}(B)$  medeixen boles i parelles a la configuració  $\bar{B}$  complementària de  $B$ . O, de manera equivalent, medeixen *forats* i parelles de forats en  $B$ . Així, per a una configuració  $B$  amb  $w$  boles ( $\bar{w} = qp - w$  forats), es compleix que

$$\bar{W}(B) = qp - W(B) \quad (2.37)$$

$$\bar{P}(B) = P(B) + (1 - q)W(B) + p \binom{q}{2}. \quad (2.38)$$

A efectes d'optimització, doncs, donat que

$$P(B) = (P|W)^{min}(w) \Leftrightarrow \bar{P}(B) = (\bar{P}|W)^{min}(w), \quad (2.39)$$

ambdues mesures són equivalents.

### 2.2.2 Mesures de les propietats d'un disseny

Tornant al cas dels dissenys en blocs, a continuació es defineixen mesures de distribució per a cada una les propietats d'un disseny amb paràmetres  $(v, b, u)$  i descriptors  $[u, r, k, \lambda]$ .

Pel que fa al nombre d'uns que, com en el cas de les boles, no depèn de la distribució, o bé els comptem directament, o bé en comptem les parelles  $P_t$  totals. Si  $U(A) = o$ , definim

$$P_t(A) = \binom{o}{2}. \quad (2.40)$$

Les cotes inferiors d'aquestes dues mesures valen

$$(U|U)^{\min}(o) = o \quad (2.41)$$

$$(P_t|U)^{\min}(o) = \binom{o}{2}, \quad (2.42)$$

i els increments són

$$\Delta(U|U)^{\min}(o) = 1 \quad (2.43)$$

$$\Delta(P_t|U)^{\min}(o) = o. \quad (2.44)$$

La uniformitat de files i columnes són dos casos particulars de distribució d'o "boles" en  $v$  i  $b$  "pots", respectivament. Les mesures corresponents són el nombre  $P_h$  de parelles *horizontals* d'uns (parelles formades amb uns pertanyents a una mateixa fila), i el nombre  $P_v$  de parelles *verticals*. Donada una configuració  $A$ , siguin  $r_i$  i  $k_j$  el nombre d'uns de la fila  $i$  i el nombre d'uns de la columna  $j$ , respectivament. Definim

$$P_h(A) = \sum_i \binom{r_i}{2} \quad (2.45)$$

$$P_v(A) = \sum_j \binom{k_j}{2}. \quad (2.46)$$

Seguint el procediment de l'apartat anterior, les seves cotes inferiors condicionades als uns són

$$(P_h|U)^{\min}(o) = v \binom{r(o)}{2} + v_0(o)r(o) \quad (2.47)$$

$$(P_v|U)^{\min}(o) = b \binom{k(o)}{2} + b_0(o)k(o), \quad (2.48)$$

on  $r(o) = \lfloor o/v \rfloor$ ,  $v_0(o) = o \bmod v$ ,  $k(o) = \lfloor o/b \rfloor$  i  $b_0(o) = o \bmod b$ , com abans. I els increments valen

$$\Delta(P_h|U)^{\min}(o) = r(o) \quad (2.49)$$

$$\Delta(P_v|U)^{\min}(o) = k(o). \quad (2.50)$$

Per a  $(v, b, o)$  tals que  $v_0(o) = 0$  o  $b_0(o) = 0$ , les cotes inferiors corresponen, respectivament, a files o columnes *estricteament* uniformes. Si a més  $o = u$ , aleshores  $r(o) = r$  i  $k(o) = k$ , els descriptors *del disseny*.

Pel que fa a l'equilibri, la correlació  $\lambda_{ik}$  entre dues files donades  $i$  i  $k$  es pot veure com el nombre de parelles verticals presents entre les dues files. Si tenim  $f = \binom{v}{2}$  parelles de files i un total  $p_v$  de parelles verticals, una configuració serà *equilibrada* quan les *parelles verticals* ("boles") estiguin uniformement distribuïdes entre les *parelles de files* ("pots"). És a dir, amb  $\lambda(p_v) = \lfloor p_v/f \rfloor$  i  $f_0(p_v) = p_v \bmod f$ , quan tinguem  $f_0(p_v)$  parelles de files amb correlació  $\lambda(p_v) + 1$  i la resta amb correlació  $\lambda(p_v)$ . Per a  $p_v$  i  $f$  tals que  $f_0(p_v) = 0$ , totes les correlacions valen exactament  $\lambda(p_v)$  i direm que la configuració és *estrictament equilibrada*.

Seguint l'analogia amb les boles, l'equilibri es pot mesurar aleshores per mitjà del nombre de parelles que es poden formar prenent parelles verticals pertanyents al mateix parell de files. Les parelles de parelles les anomenarem *quàdruples* i el seu nombre en una configuració  $A$  ve donat per

$$Q(A) = \sum_{i=1}^{v-1} \sum_{k=i+1}^v \binom{\lambda_{ik}}{2}. \quad (2.51)$$

La seva cota inferior condicionada a les *parelles verticals* és, aleshores,

$$(Q|P_v)^{\min}(p_v) = f \binom{\lambda(p_v)}{2} + f_0(p_v)\lambda(p_v), \quad (2.52)$$

amb increment

$$\Delta(Q|P_v)^{\min}(p_v) = \lambda(p_v). \quad (2.53)$$

Amb  $p_v < f$  la condició d'equilibri correspon a  $\lambda(p_v) = 0$  i  $Q(A) = 0$ . Per a  $p_v = f$ , l'equilibri s'assoleix amb  $\lambda(p_v) = 1$  i  $f_0(p_v) = 0$ , i  $Q$  encara val 0. I per a  $p_v > f$ , tenim  $\lambda(p_v) \geq 1$  i  $Q > 0$ . De fet, per a  $p_v \geq f$ ,  $(Q|P_v)^{\min}(p_v)$  és *monòtonament creixent*.

La condició d'equilibri no implica columnes uniformes (figura 1.1). El mateix nombre  $p_v$  de parelles verticals es pot assolir amb força uns ben distribuïts entre columnes, o amb menys uns més mal distribuïts. I, viceversa, per a un nombre fixat  $o$  d'uns, el nombre de parelles verticals depèn de la seva distribució per columnes. Donat  $o$ , doncs, els valors  $\lambda(p_v)$  i  $f_0(p_v)$  no queden unívocament determinats. Sota hipòtesi de columnes uniformes, però, el nombre de parelles verticals es minimitza (eq. 2.48), i el denotarem per  $p_v^o$ . En aquestes circumstàncies,  $\lambda(p_v^o)$  i  $f_0(p_v^o)$  són fixos, i la cota inferior de les quàdruples  $(Q|U)^{\min}(o)$ , *condicionada al nombre d'uns*, ve donada per

$$(Q|U)^{\min}(o) = (Q|P_v)^{\min}(p_v^o) = f \binom{\lambda(p_v^o)}{2} + f_0(p_v^o)\lambda(p_v^o), \quad (2.54)$$

amb increment

$$\Delta(Q|U)^{\min}(o) = k(o)\lambda(p_v^o) + f_1(p_v^o), \quad (2.55)$$

on  $f_1(p_v^o) = \max\{0, k(o) - f + f_0(p_v^o)\}$ . L'expressió de l'increment es justifica com segueix. Pel fet d'afegir un 1 la cota inferior de les parelles verticals s'incrementa en

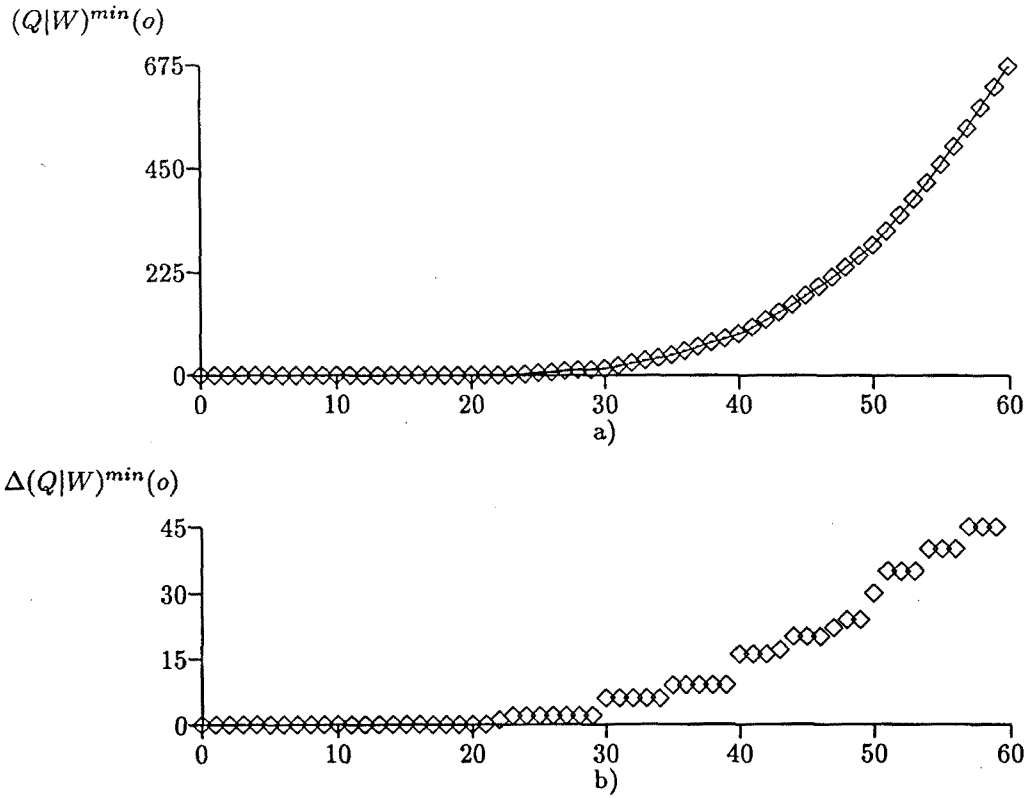


Figura 2.3: a) Cota inferior de  $Q(A)$  en funció del nombre d'uns i b) increment per al cas  $v = 6$ ,  $b = 10$ .

$k(o)$ . Si el nombre  $f - f_0(p_v^o)$  de parelles de files que correlaven exactament  $\lambda(p_v^o)$  és suficient (és menor que  $k(o)$ ), aleshores cada una de les  $k(o)$  parelles verticals contribueix en  $\lambda(p_v^o)$  quàdruples a la nova cota inferior de  $Q$ . Però les parelles verticals que sobrin ( $k(o) - f + f_0(p_v^o)$  parelles), contribueixen en  $\lambda(p_v^o) + 1$ . La figura 2.3 mostra les expressions anteriors en funció d' $o$ , per al cas  $v = 6$  i  $b = 10$ . Les expressions anteriors descriuen una configuració amb  $o$  uns que tingui columnes uniformes i nombre de quàdruples mínim. Si a més  $o = u$ , aleshores  $f_0(p_v^o) = 0$  i  $\lambda(p_v^o) = \lambda$ , el descriptor del disseny.

La discussió anterior demostra, per construcció, que si  $A$  és una configuració equilibrada amb  $o$  uns i té columnes uniformes, aleshores  $Q(A) = (Q|U)^{min}(o)$ . L'afirmació inversa només és certa en el rang en que  $(Q|P_v)^{min}(p_v)$  és monòtonament creixent. És a dir, quan  $o$  és tal que  $p_v \geq f$ . Formalment,

**Proposició 2.3** *Sigui*

$$o_1 = [(P_v|U)^{min}]^{-1}(f). \quad (2.56)$$

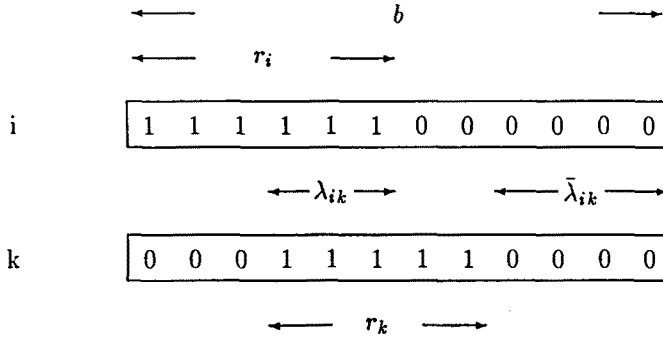


Figura 2.4: Correlació  $\bar{\lambda}_{ik}$  entre els zeros de dos files  $i$  i  $k$  que correnen  $\lambda_{ik}$  i tenen  $r_i$  i  $r_k$  uns, respectivament.

Donada una configuració  $A - v \times b$  tal que  $U(A) = o \geq o_1$

$$Q(A) = (Q|U)^{\min}(o) \Leftrightarrow A \text{ és equilibrada i té columnes uniformes.} \quad (2.57)$$

Aquest és un resultat important que condueix, més endavant, al teorema 2.4.

L'analogia amb les boles s'acaba quan ens referim al nombre de quàdruples  $\bar{Q}(A)$  de la configuració complementària. En aquest cas, si per a un parell de files  $(i, k)$  la correlació en  $A$  val  $\lambda_{ik}$ , la correlació en  $\bar{A}$  val

$$\bar{\lambda}_{ik} = b + \lambda_{ik} - r_i - r_k, \quad (2.58)$$

tal com s'il·lustra a la figura 2.4. El terme "complementari" fa referència en aquest cas als uns i zeros de la matriu  $A$ , i no es correspon, com ho feia en les boles, al complementari d'un "pot" (parell de files) de capacitat  $b$  amb  $\lambda$  "boles" (parelles verticals), ja que li correspondrien  $b - \lambda$  "boles". Així, l'equació (2.38) no es verifica i  $\bar{Q}(A)$  no té una expressió senzilla en termes de  $Q(A)$  i  $P_v(A)$ . Per calcular-ne la cota inferior i l'increment, doncs, caldrà recórrer directament a les expressions (2.10) i (2.12). En el cas general doncs, donats  $(v, b, o)$ , la minimització de  $Q$  no implica la minimització de  $\bar{Q}$ , i farem ús explícit de les dues mesures.

## 2.3 Optimalitat

Les mesures que acabem de definir són  $U(A)$ ,  $P_t(A)$ ,  $P_h(A)$ ,  $P_v(A)$ ,  $Q(A)$  i  $\bar{Q}(A)$ . En aquest apartat veurem quin valor prenen sota la hipòtesi d'optimalitat (en un disseny hipotètic), i en quines circumstàncies podem afirmar que una configuració és un disseny (condicions necessàries i suficients d'optimalitat).



$M$	$M^*$	$\Delta^- M^*$	$\Delta^+ M^*$
$U$	$u$	1	1
$P_t$	$\binom{u}{2}$	$u - 1$	$u$
$P_h$	$v \binom{r}{2}$	$r - 1$	$r$
$P_v$	$b \binom{k}{2}$	$k - 1$	$k$
$Q$	$f \binom{\lambda}{2}$	$(k - 1)(\lambda - 1)$	$k\lambda$
$\bar{Q}$	$f \binom{\lambda}{2}$	$-\bar{k}\bar{\lambda}$	$-(\bar{k} - 1)(\bar{\lambda} - 1)$

Taula 2.1: Donat un conjunt  $(v, b, u)$  de paràmetres admissible, valors i increments òptims de les mesures  $U$ ,  $P_t$ ,  $P_h$ ,  $P_v$ ,  $Q$  i  $\bar{Q}$ , en termes dels descriptors del disseny (i del seu complementari).

### 2.3.1 Valors òptims

Donat un conjunt  $(v, b, u)$  de paràmetres admissible, amb descriptors  $[u, r, k, \lambda]$ , la *hipòtesi d'optimalitat* consisteix en assumir que les propietats i) a iv) de la definició 1.2 de disseny es satisfan. L'aplicació d'aquesta hipòtesi a l'*avaluació* de les mesures anteriors, però, no implica l'existència d'una configuració que les satisfaci ja que, com hem dit, el disseny pot no existir.

Definim el *valor òptim*  $M^*$  d'una mesura  $M(A)$  com el valor que prendria sota la hipòtesi d'optimalitat. Donat que, per la proposició 2.1, les mesures que estem estudiant són mesures de distribució, sota la hipòtesi d'optimalitat cada una d'elles assoleix la seva cota inferior. Per a  $M \in \{U, P_t, P_h, Q, \bar{Q}\}$  es compleix doncs que

$$M^* = (M|U)^{\min}(u). \quad (2.59)$$

Anàlogament, definim els *increments òptims* d'una mesura com els increments de la seva cota inferior avaluats a  $o = u - 1$  i  $o = u$ , respectivament, i els denotarem per

$$\Delta^- M^* \equiv \Delta(M|U)^{\min}(u - 1) \quad (2.60)$$

$$\Delta^+ M^* \equiv \Delta(M|U)^{\min}(u). \quad (2.61)$$

La taula 2.1 mostra aquests valors i increments òptims, expressats en termes dels descriptors del disseny.

A continuació, definim l'*increment local òptim*  $\Delta^{ij} M^*$  d'una mesura respecte d'un component  $(i, j)$  com el valor que prendria  $\Delta^{ij} M$  sota la hipòtesi d'optimalitat, de manera que medeix la diferència entre un disseny hipotètic i una configuració *veïna*. La figura 2.5 il·lustra el càlcul de l'increment local en el cas més complicat: les quàdruples. Permutant files i columnes per tal d'agrupar els 1s de la fila  $i$  i els 1s de la

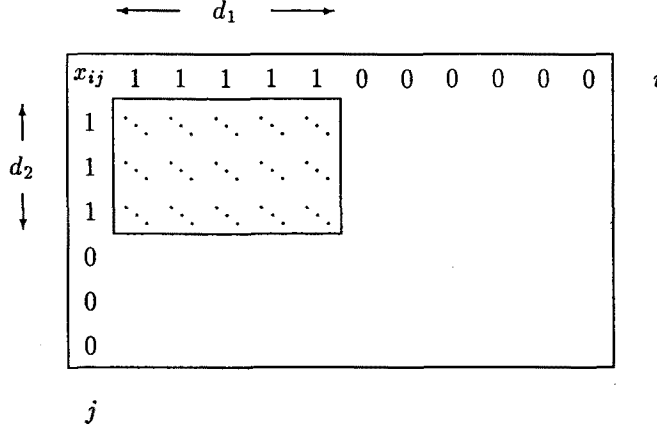


Figura 2.5: L'increment local  $\Delta^{ij}Q$  és igual al nombre d'uns dins l'àrea sombreada (veure text).

columna  $j$  tal com es mostra a la figura, l'increment  $\Delta^{ij}Q$  en el nombre de quàdruples derivat de commutar  $x_{ij}$  és igual al nombre d'uns dins l'àrea sombreada. Sota la hipòtesi d'optimalitat, files columnes i correlacions són estrictament uniformes de manera que, amb independència de la posició  $(i, j)$ , si  $x_{ij} = 1$  les dimensions  $d_1 \times d_2$  de l'àrea sombreada són  $(r - 1) \times (k - 1)$ , i si  $x_{ij} = 0$ , són  $r \times k$ . Donat que sabem que cada una de les files afectada correla  $\lambda$  amb la fila  $i$ , aleshores  $\Delta^{ij}Q^*|_{x_{ij}=1} = (k - 1)(\lambda - 1)$ , i  $\Delta^{ij}Q^*|_{x_{ij}=0} = k\lambda$ . Una situació similar es produeix per a la resta de mesures.

Generalitzant, doncs, la hipòtesi d'optimalitat redueix el repertori de valors dels increments locals a dos únics valors, de manera que definim

$$\Delta^1 M^* \equiv \Delta^{ij} M^*|_{x_{ij}=1} \tag{2.62}$$

$$\Delta^0 M^* \equiv \Delta^{ij} M^*|_{x_{ij}=0}. \tag{2.63}$$

Si avaluem aquests valors, i els comparem amb la taula 2.1, es verifica que

$$\Delta^1 M^* = \Delta^- M^* \tag{2.64}$$

$$\Delta^0 M^* = \Delta^+ M^*. \tag{2.65}$$

Així, *qualsevol* configuració veïna d'una òptima és màximament distribuïda en el sentit de les propietats considerades. En el capítol 6 parlarem en més detall d'aquesta propietat dels dissenys en blocs.

### 2.3.2 Condicions necessàries i suficients d'optimalitat

Per a  $M \in \{U, P_t, P_v, P_h, Q, \bar{Q}\}$ , les expressions (2.59) i (2.64, 2.65) són *condicions necessàries d'optimalitat* en el sentit que, si existeix una configuració  $A^*$  corresponent a un disseny en blocs, aleshores  $M(A^*) = M^*$ , i  $\Delta^{ij}M(A^*)$  val  $\Delta^1M^*$  quan  $x_{ij} = 1$  i val  $\Delta^0M^*$  quan  $x_{ij} = 0$ .

Formulat a l'inrevés, donada una configuració  $A$ ,  $M(A) = M^*$  per si sol no implica res (poden haver-hi pocs uns i estar mal distribuïts). Però si  $U(A) = u$ , aleshores  $M(A) = M^*$  implica màxima distribució de la propietat corresponent. Així, i en el supòsit del nombre d'uns just, per la proposició 2.1  $P_h(A) = P_h^*$  implica files uniformes,  $P_v(A) = P_v^*$  implica columnes uniformes, i  $Q(A) = Q^*$  implica equilibri en les correlacions. A més, la definició 1.1 restringeix els paràmetres  $(v, b, u)$  admissibles als casos en que el descriptor d'equilibri  $\lambda > 0$ , de manera que s'ha de complir  $u \geq o_1$  (eq. 2.56). Així doncs, per la proposició 2.3,  $Q(A) = Q^*$  implica també columnes uniformes.

Equilibri i columnes uniformes, però, defineixen un disseny (definició 1.1), de manera que, resumint

**Teorema 2.4** *Donada l'admissibilitat dels paràmetres  $(v, b, u)$ , una configuració  $A \in \mathcal{A} - v \times b$  és la matriu d'incidència d'un disseny en blocs  $DB(v, b, u)$  si i només si*

$$U(A) = u \quad i \quad (2.66)$$

$$Q(A) = Q^*. \quad (2.67)$$

Expandides, aquestes dues expressions constitueixen la definició més breu de disseny en blocs, en termes de les variables d'estat  $x_{ij}$ , i la resta de propietats s'en deriven.

A continuació es presenten una sèrie de formulacions alternatives del teorema anterior, que es faran servir al llarg del treball.

**Corol·lari 2.4.1** *Donada l'admissibilitat dels paràmetres  $(v, b, u)$ , cada una de les parelles d'expressions que es formulen a continuació són condició necessària i suficient per a que una configuració  $A \in \mathcal{A} - v \times b$  sigui la matriu d'incidència d'un disseny en blocs  $DB(v, b, u)$ .*

i

$$U(A) = u, \quad (2.68)$$

$$\lambda_{ik} = \lambda \quad \forall (i, k). \quad (2.69)$$

ii

$$Q(A) = Q^*, \quad (2.70)$$

$$\bar{Q}(A) = \bar{Q}^*. \quad (2.71)$$

iii

$$U(A) = u, \quad (2.72)$$

$$\Delta^{ij}Q(A) = \Delta^1Q^* \quad \forall(i,j) : x_{ij} = 1. \quad (2.73)$$

iv

$$\Delta^{ij}Q(A) = \Delta^1Q^* \quad \forall(i,j) : x_{ij} = 1, \quad (2.74)$$

$$\Delta^{ij}Q(A) = \Delta^0Q^* \quad \forall(i,j) : x_{ij} = 0. \quad (2.75)$$

v

$$\Delta^{ij}Q(A) = \Delta^1Q^* \quad \forall(i,j) : x_{ij} = 1, \quad (2.76)$$

$$\Delta^{ij}\bar{Q}(A) = \Delta^1\bar{Q}^* \quad \forall(i,j) : \bar{x}_{ij} = 1. \quad (2.77)$$

Les condicions necessàries es deriven directament de la definició de valors óptims (calculats a l'apartat anterior sota la hipòtesi d'optimalitat). Pel que fa a les condicions suficients,

- i Aplicant  $\lambda_{ik} = \lambda$  a l'equació (2.51) obtenim  $Q = Q^*$ , que ens remet al teorema.
- ii  $Q = Q^* = (Q|U)^{min}(u)$  implica  $o \leq u$  ( $u$  és el màxim nombre d'uns que en distribució òptima no supera les  $Q^*$  quàdruples). I la mateixa afirmació en el cas complementari implica  $\bar{o} \leq \bar{u}$ . De manera que  $o = u$ , que ens remet novament al teorema.
- iii Pel que fa als increments locals, el nombre total de quàdruples es pot expressar com

$$Q(A) = \frac{1}{4} \sum_{i=1}^v \sum_{j=1}^b x_{ij} \Delta^{ij}Q(A). \quad (2.78)$$

Si es compleixen les hipòtesis del corol·lari,

$$Q(A) = \frac{1}{4}u\Delta^1Q^* = \frac{1}{4}u(k-1)(\lambda-1)$$

que, aplicant les condicions d'admissibilitat, es pot reescriure com

$$Q(A) = \frac{v(v-1)}{2} \frac{\lambda(\lambda-1)}{2} = \binom{v}{2} \binom{\lambda}{2},$$

que és l'expressió de  $Q^*$ .

- iv Procedint com en el cas anterior, si el nombre d'uns de la configuració  $A$  és  $o$ , aleshores

$$Q(A) = \frac{1}{4}o\Delta^1Q^*.$$

Si commutem un dels 1s, la configuració resultant  $A^-$  tindrà  $Q(A^-) = Q(A) - \Delta^1 Q^*$  quàdruples. Ara bé, avaluant aquesta expressió, i comparant-la amb la cota inferior, es verifica que, si  $o < u$  aleshores  $Q(A^-) < (Q|U)^{\min}(o - 1)$ , en contradicció amb la definició de la cota. D'on es deriva que  $o \geq u$ . De manera similar, si commutem un dels zeros la configuració resultant  $A^+$  tindrà  $Q(A^+) = Q(A) + \Delta^0 Q^*$ . Però si  $o > u$  aleshores  $Q(A^+) < (Q|U)^{\min}(o + 1)$ . D'on  $o \leq u$  i, amb el resultat anterior,  $o = u$  i  $Q(A) = Q^*$ .

- v La primera part de la demostració anterior diu:  $\forall(i, j) : x_{ij} = 1$ , si  $\Delta^{ij} Q(A) = \Delta^1 Q^*$  aleshores  $o \geq u$ . Aplicant el mateix argument a la configuració complementària, la hipòtesi (2.77) implica  $\bar{o} \geq \bar{u}$ . De manera que, com abans,  $o = u$  i  $Q(A) = Q^*$ .

## 2.4 Funcions de cost per a la generació de dissenys

Una mesura  $F$  és una funció de cost per a la generació de dissenys en blocs si totes les configuracions que són dissenys prenen el valor òptim, i totes les que no ho són prenen un valor més gran. En aquest apartat es defineix un repertori de funcions de cost basades en les mesures de les propietats d'un disseny.

### 2.4.1 Definició i forma normal

**Definició 2.5** Donat un conjunt de paràmetres admissible  $(v, b, u)$ , direm que  $F : \mathcal{A} \rightarrow \mathcal{R}$  és una funció de cost per a la generació de dissenys en blocs si i només si compleix les condicions següents:

- i) Existeix una cota  $F^*$  tal que  $F(A) \geq F^*$ ,  $\forall A \in \mathcal{A}$ .
- ii) Si existeix un disseny  $DB(v, b, u)$  amb matriu d'incidència  $A^*$ , aleshores  $F(A^*) = F^*$ .
- iii) Si existeix una configuració  $A^*$  tal que  $F(A^*) = F^*$ , aleshores  $A^*$  és la matriu d'incidència d'un disseny  $DB(v, b, u)$ .

Donat que una funció de cost ha de medir les propietats d'un disseny i, eventualment, les del disseny complementari, l'expressió genèrica que es proposa en aquest treball és la combinació lineal de les mesures dels apartats anteriors. Suprimint, per breujar, les dependències en  $A$  tenim

$$F = \rho_u U + \rho_t P_t + \rho_h P_h + \rho_v P_v + \rho_q Q + \rho_{\bar{q}} \bar{Q}, \quad (2.79)$$

amb  $\rho_t, \rho_h, \rho_v, \rho_q$ , i  $\rho_{\bar{q}}$  no negatius, ja que es tracta de minimitzar ( $\rho_u$ , però, és lliure). Aquesta expressió defineix a partir d'ara la *forma normal* d'una funció de cost, que queda unívocament definida pel valor dels seus *coeficients*

$$F \equiv (\rho_u, \rho_t, \rho_h, \rho_v, \rho_q, \rho_{\bar{q}}). \quad (2.80)$$

## 2.4.2 Condicions de validesa d'una funció de cost

L'expressió (2.79) és una mesura genèrica de la distribució dels paràmetres d'un disseny. Per fer-ne una funció de cost cal, en primer lloc, que els termes que hi intervenen amb coeficients no nuls siguin suficients per garantir l'optimalitat. I cal, en segon lloc, que la cota inferior  $(F|U)^{\min}(o)$  tingui un *mínim puntual* a  $o = u$ .

Seguint el procediment de la proposició 2.2, el segon objectiu és factible ja que l'admissibilitat dels paràmetres garanteix que es satisfan totes les condicions de multiplicitat. Les cotes inferiors dels tres tipus de parelles i els dos tipus de quàdruples són globalment convexes (increments no decreixents), el que garanteix que el mínim sigui també global i, com a la proposició esmentada, el mínim puntual es pot ajustar fixant el pas per zero de  $\Delta(F|U)^{\min}(o)$  a  $o = u$ . Aconseguit això, i pel teorema 2.4, l'única restricció addicional per satisfer el primer objectiu és que la funció sigui sensible a les quàdruples. Formalment, doncs,

**Teorema 2.5** *Donada l'admissibilitat dels paràmetres  $(v, b, u)$ , la mesura  $F$  de l'equació (2.79) és una funció de cost per a la generació de dissenys en blocs si*

$$\rho_q > 0 \quad (2.81)$$

i

$$\Delta^- F^* = \Delta(F|U)^{\min}(u-1) < 0 \quad (2.82)$$

$$\Delta^+ F^* = \Delta(F|U)^{\min}(u) > 0. \quad (2.83)$$

*El seu valor òptim és*

$$\begin{aligned} F^* &= (F|U)^{\min}(u) = \\ &= \rho_u U^* + \rho_t P_t^* + \rho_h P_h^* + \rho_v P_v^* + \rho_q Q^* + \rho_{\bar{q}} \bar{Q}^*. \end{aligned} \quad (2.84)$$

Fixat  $F^*$ , les condicions d'optimalitat de l'apartat 2.3.2 garanteixen que totes i únicament les configuracions corresponents a dissenys en blocs assoleixen el mínim global d' $F$ . De manera que  $F$  és una funció de cost (definició 2.5).

Les expressions (2.82, 2.83) s'avaluen en termes dels increments òptims de cada component (taula 2.1), i defineixen el marge de valors permesos als coeficients.

### 2.4.3 Elecció d'un repertori

Definim l'*estructura* d'una funció de cost com el conjunt de termes que hi intervenen amb coeficients no nuls, i la denotarem per mitjà de subíndexs. Així, per exemple,  $F_{uvq}$  fa referència a qualsevol funció amb  $\rho_t = \rho_h = \rho_{\bar{q}} = 0$  i  $\rho_u, \rho_v, \rho_q \neq 0$ . Fixada l'estructura, la *composició* d'una funció de cost particular descriu el pes relatiu de cada terme, i ve donada pel valor dels seus coeficients. Donat que l'optimització no ve afectada per l'escala, utilitzarem sempre coeficients enters i, malgrat la redundància, escriurem els 6 valors. Així  $F_{uvq}(-4, 0, 0, 1, 2, 0)$  és un cas particular de l'exemple anterior.

Dins les restriccions fixades a l'apartat anterior, i a fi de reduir el repertori a dimensions abastables, en aquest apartat es defineixen dos criteris per a la selecció d'una funció de cost. En primer lloc, en termes estructurals, tractarem per separat les funcions amb  $\rho_{\bar{q}} = 0$  i les funcions amb  $\rho_{\bar{q}} \neq 0$ . En el primer grup, les estructures possibles són

$$F_{uq} \quad (2.85)$$

$$F_{utq} \quad (2.86)$$

$$F_{uhq} \quad (2.87)$$

$$F_{uvq} \quad (2.88)$$

$$F_{uthq} \quad (2.89)$$

$$F_{utvq} \quad (2.90)$$

$$F_{uhvq} \quad (2.91)$$

$$F_{uthvq} \quad (2.92)$$

El terme  $Q$  és imprescindible, pel teorema 2.4, i el terme  $U$  també ho és, ja que és el que ens permetrà fixar el mínim. Del segon grup, l'única estructura que considerarem és

$$F_{q\bar{q}}, \quad (2.93)$$

que es deriva del corol·lari 2.4.1 ii), i prescindirem de les combinacions mixtes.

En segon lloc, i en termes de composició, ens cenyirem únicament a funcions de cost *simètriques* (eq. 2.33),

$$\Delta^- F^* = -\Delta^+ F^*, \quad (2.94)$$

que igualen la magnitud dels increments óptims, amb l'objectiu d'equilibrar l'estabilitat dels uns i dels zeros.

Per a les estructures del primer grup, la composició relativa dels termes  $M_i \in \{P_t, P_h, P_v, Q\}$  s'escollirà experimentalment (capítol 5) i, per la condició de simetria,

$$\alpha_u = - \sum_i \rho_i \frac{\Delta^- M_i^* + \Delta^+ M_i^*}{2}, \quad (2.95)$$

amb  $\rho_i$  el coeficient corresponent a cada terme, i  $\alpha_u$  el valor de  $\rho_u$  que satisfà la condició ( $\alpha_u$  és, en conseqüència, paramètric en els coeficients de la funció i els

paràmetres del disseny). L'expressió  $F_{uvq}(\alpha_u, 0, 0, 1, 2, 0)$ , per exemple, denotarà aleshores la corresponent funció de cost *simètrica*.

En el cas de l'estructura  $F_{uq}$ , l'únic terme que intervé a 2.95 és  $Q$ . De manera que la seva composició és fixa, i la funció corresponent,  $F_{uq}(\alpha_u, 0, 0, 0, 1, 0)$ , es denotarà simplement com  $F_{uq}$ .

Procedint anàlogament, la condició de simetria aplicada a l'estructura  $F_{q\bar{q}}$  del segon grup condueix a una única funció amb

$$\alpha_q = -\frac{\Delta^- \bar{Q}^* + \Delta^+ \bar{Q}^*}{2} \quad (2.96)$$

$$\alpha_{\bar{q}} = \frac{\Delta^- Q^* + \Delta^+ Q^*}{2}, \quad (2.97)$$

que denotarem per  $F_{q\bar{q}}(0, 0, 0, 0, \alpha_q, \alpha_{\bar{q}})$ , o, simplement,  $F_{q\bar{q}}$ .

## 2.5 La xarxa optimitzadora

A l'hora de definir una xarxa (apartat 1.2), el procediment habitual consisteix en escollir una arquitectura estructuralment equivalent a la funció de cost que es vol realitzar. L'estat de les unitats codifica aleshores l'estat o *configuració* actual. Les connexions representen les restriccions entre les variables d'estat, i l'energia de la xarxa representa el cost de la configuració.

A fi d'obtenir una xarxa genèricament equivalent a les funcions de cost de l'apartat anterior, prenem  $v \times b$  unitats organitzades en  $v$  files i  $b$  columnes, de manera que l'estat  $x_{ij}$  de la unitat  $(i, j)$  representi l'entrada corresponent de la matriu d'incidència. A continuació definim les connexions següents:

**Llindars:** Auto-connexions, definides sobre  $\{(i, j)\}$ . Direm que l'auto-connexió a  $(i, j)$  està *activa* quan  $x_{ij} = 1$ . I el nombre total d'auto-connexions actives en una configuració  $A$  ve donat per

$$N_u \equiv \sum_{i=1}^v \sum_{j=1}^b x_{ij}, \quad (2.98)$$

que és el nombre total d'uns  $U(A)$ . Totes les auto-connexions tenen el mateix *pes*, que denotarem per  $\nu_u$ .

**Parelles totals:** Connexions definides sobre *qualsevol* parella d'unitats  $\{(i, j), (k, l)\} : (i, j) \neq (k, l)\}$ . Una parella està *activa* quan les dues unitats que la formen valen 1, i el nombre total de parelles actives en una configuració  $A$  és

$$N_t = \sum_{i=1}^v \sum_{k=i}^v \sum_{j=1}^b \sum_{l=j}^b x_{ij} x_{kl} - N_u, \quad (2.99)$$

que és el nombre total de parelles d'uns  $P_t(A)$ . El pes serà el mateix per a totes les parelles, i el denotarem per  $\nu_t$ .



**Connexions horitzontals:** Connexions definides sobre les parelles d'unitats  $\{[(i, j), (i, l)] : l \neq j\}$ , és a dir connexions entre parelles d'unitats disposades en horitzontal (a la mateixa fila). Una connexió horitzontal està *activa* quan les dues unitats que la formen valen 1, i el nombre total de connexions horitzontals actives en una configuració  $A$  és

$$N_h = \sum_{i=1}^v \sum_{j=1}^{b-1} \sum_{l=j+1}^b x_{ij} x_{il}, \quad (2.100)$$

que és el nombre de parelles horitzontals d'uns,  $P_h(A)$ . El pes serà el mateix per a totes les connexions horitzontals, i el denotarem per  $\nu_h$ .

**Connexions verticals:** Connexions

definides sobre les parelles d'unitats  $\{[(i, j), (k, j)] : k \neq i\}$ , és a dir connexions entre parelles d'unitats disposades en vertical (a la mateixa columna). Una connexió vertical està *activa* quan les dues unitats que la formen valen 1, i el nombre total de connexions verticals actives en una configuració  $A$  és

$$N_v = \sum_{j=1}^b \sum_{i=1}^{v-1} \sum_{k=i+1}^v x_{ij} x_{kj}, \quad (2.101)$$

que és el nombre total de parelles verticals d'uns,  $P_v(A)$ . El pes serà el mateix per a totes les connexions verticals, i el denotarem per  $\nu_v$ .

**Quàdruples d'uns:** Connexions definides sobre quàdruples d'unitats  $\{[(i, j), (i, l), (k, j), (k, l)] : k \neq i, l \neq j\}$ . Una quàdrupla d'uns està *activa* quan les quatre unitats que la formen valen 1, i el nombre total de quàdruples d'uns actives en una configuració  $A$  és

$$N_q = \sum_{i=1}^{v-1} \sum_{k=i+1}^v \sum_{j=1}^{b-1} \sum_{l=j+1}^b x_{ij} x_{il} x_{kj} x_{kl}, \quad (2.102)$$

que correspon a  $Q(A)$ . El pes serà el mateix per a totes les quàdruples d'uns, i el denotarem per  $\nu_q$ .

**Quàdruples de zeros:** Connexions definides sobre quàdruples d'unitats  $\{[(i, j), (i, l), (k, j), (k, l)] : k \neq i, l \neq j\}$ . Una quàdrupla de zeros està *activa* quan les quatre unitats que la formen valen 0, i el nombre total de quàdruples zeros actives en una configuració  $A$  és

$$N_{\bar{q}} = \sum_{i=1}^{v-1} \sum_{k=i+1}^v \sum_{j=1}^{b-1} \sum_{l=j+1}^b \bar{x}_{ij} \bar{x}_{il} \bar{x}_{kj} \bar{x}_{kl}, \quad (2.103)$$

que correspon a  $\bar{Q}(A)$ . El pes serà el mateix per a totes les quàdruples de zeros, i el denotarem per  $\nu_{\bar{q}}$ .

Donada la presència de les quadruples, l'estructura proposada és una xarxa d'ordre superior (apartat 1.2.3).

L'*energia* de la xarxa es defineix com la suma dels pesos de les connexions actives,

$$E = \nu_u N_u + \nu_t N_t + \nu_h N_h + \nu_v N_v + \nu_q N_q + \nu_{\bar{q}} N_{\bar{q}}. \quad (2.104)$$

Si, per a una funció de cost concreta  $F(\rho_u, \rho_t, \rho_h, \rho_v, \rho_q, \rho_{\bar{q}})$  fem

$$\nu_i = \rho_i, \forall i \in \{u, t, v, q, \bar{q}\}, \quad (2.105)$$

aleshores, comparant l'expressió anterior amb l'equació (2.79), i en virtut de les equivalències que acabem de formular, obtenim

$$E(A) = F(A). \quad (2.106)$$

Per a qualsevol configuració  $A$ , l'energia de la xarxa *avalua* el valor corresponent de la funció de cost. Es diu aleshores que l'equació (2.105) *projecta* la funció de cost sobre l'energia de la xarxa, i l'òptim es produeix si, i només si,

$$E(A) = E^* = F^*, \quad (2.107)$$

amb  $F^*$  donat per l'equació (2.84).

El *camp local*  $\phi_{ij}$  es defineix com la suma ponderada de les connexions actives que arriben a la unitat  $(i, j)$ . Per a xarxes d'ordre superior (eq. 1.28), i aplicat al nostre cas,

$$\phi_{ij} = \nu_u \Delta^{ij} N_u + \nu_t \Delta^{ij} N_t + \nu_h \Delta^{ij} N_h + \nu_v \Delta^{ij} N_v + \nu_q \Delta^{ij} N_q + \nu_{\bar{q}} \Delta^{ij} N_{\bar{q}}, \quad (2.108)$$

amb

$$\Delta^{ij} N_u = 1 \quad (2.109)$$

$$\Delta^{ij} N_t = \left( \sum_k \sum_l x_{kl} \right) - x_{ij} \quad (2.110)$$

$$\Delta^{ij} N_h = \sum_{l \neq j} x_{il} \quad (2.111)$$

$$\Delta^{ij} N_v = \sum_{k \neq i} x_{kj} \quad (2.112)$$

$$\Delta^{ij} N_q = \sum_{k \neq i} \sum_{l \neq j} x_{il} x_{kj} x_{kl} \quad (2.113)$$

$$\Delta^{ij} N_{\bar{q}} = \sum_{k \neq i} \sum_{l \neq j} \bar{x}_{il} \bar{x}_{kj} \bar{x}_{kl}, \quad (2.114)$$

d'on es deriva que el camp local correspon a l'increment local d'energia respecte a la unitat  $(i, j)$ ,

$$\phi_{ij} = \Delta^{ij} E. \quad (2.115)$$

Les expressions anteriors són equivalents a  $\Delta^{ij}U$ ,  $\Delta^{ij}P_t$ ,  $\Delta^{ij}P_h$ ,  $\Delta^{ij}P_v$ ,  $\Delta^{ij}Q$  i  $\Delta^{ij}\bar{Q}$ , respectivament, i els camps locals òptims queden consegüentment determinats (eqs. 2.64, 2.65, i taula 2.1). El camp local medeix, aleshores, l'increment local de la funció  $F$ ,

$$\phi_{ij} = \Delta^{ij}F, \quad (2.116)$$

i el procés de relaxació de la xarxa equival a la *minimització* de la funció.

Com es pot verificar, per a cada tipus de connexió, la suma dels camps locals de les unitats actives, dividida per l'aritat, medeix el nombre total de connexions actives d'aquell tipus

$$N_u = \sum_i \sum_j x_{ij} \Delta^{ij} N_u \quad (2.117)$$

$$N_t = \frac{1}{2} \sum_i \sum_j x_{ij} \Delta^{ij} N_t \quad (2.118)$$

$$N_h = \frac{1}{2} \sum_i \sum_j x_{ij} \Delta^{ij} N_h \quad (2.119)$$

$$N_v = \frac{1}{2} \sum_i \sum_j x_{ij} \Delta^{ij} N_v \quad (2.120)$$

$$N_q = \frac{1}{4} \sum_i \sum_j x_{ij} \Delta^{ij} N_q \quad (2.121)$$

$$N_{\bar{q}} = \frac{1}{4} \sum_i \sum_j x_{ij} \Delta^{ij} N_{\bar{q}}. \quad (2.122)$$

Computacionalment, les xarxes optimitzadores són especialment adequades per a la cerca local doncs, com es desprèn de les expressions anteriors, la complexitat de l'avaluació dels increments és un ordre inferior a l'avaluació explícita de la funció de cost.

## 2.6 Funcions de cost de tipus $H$ . Error quadràtic dels camps locals

En la generació de dissenys, el fet de conèixer les propietats de l'òptim permet reconèixer amb facilitat quan una configuració és un disseny (condicions d'optimalitat). Però a més, coneixent els increments òptims, podem medir també si n'està molt allunyada. En aquest apartat es proposa la suma d'errors quadràtics entre els increments locals actual i òptim com a mesura de la distància entre una configuració qualsevol i un disseny. I en conseqüència, com a funció de cost per a la generació de dissenys [Bofill & Torras, 93a, Fontdecaba, 94].

Com veïem a l'apartat 2.3.1, per a qualsevol funció de cost  $F$ , el valor òptim de l'increment local respecte a la unitat  $(i, j)$  és  $\Delta^1 F^*$  quan  $x_{ij} = 1$ , i  $\Delta^0 F^*$  quan

$x_{ij} = 0$ . Per a una configuració  $A$  donada, si definim l'error local  $\varepsilon_{ij}$  com

$$\varepsilon_{ij} = \begin{cases} \Delta^{ij}F(A) - \Delta^1F^* & \text{quan } x_{ij} = 1 \\ \Delta^{ij}F(A) - \Delta^0F^* & \text{quan } x_{ij} = 0, \end{cases} \quad (2.123)$$

la funció

$$H(A) = \sum_{i=1}^v \sum_{j=1}^b \varepsilon_{ij}^2 \quad (2.124)$$

medeix l'error quadràtic global  $H(A)$  entre els increments locals de la configuració  $A$  i els d'un disseny hipotètic. En aquests termes,

**Proposició 2.6** *Donada una funció de cost  $F$  que satisfaci el teorema 2.5, l'error quadràtic  $H$  de l'equació (2.124) és també una funció de cost vàlida per a la generació de dissenys, i el seu valor òptim és*

$$H^* = 0. \quad (2.125)$$

El valor òptim  $H(A) = H^*$  s'assoleix si i només si tots els increments locals són òptims i, pel corol·lari 2.4.1 iv), això pot succeir si i només si  $A$  correspon a un disseny.

L'estructura d'una funció de tipus  $H$  és molt diferent de l'estructura de la corresponent funció  $F$ , i no es pot projectar directament sobre una xarxa optimitzadora convencional (una xarxa amb pesos fixes i unitats variables). Malgrat això, si entenem el camp local com l'increment local de la funció respecte a una unitat  $(i, j)$  (eq. 2.116), el camp local equivalent de la funció  $H$  ve definit per

$$\phi_{ij}^H \equiv \Delta^{ij}H = H|_{x_{ij}=1} - H|_{x_{ij}=0}. \quad (2.126)$$

Aplicat a la definició d' $H$ ,  $\phi_{ij}^H$  s'expressa aleshores com

$$\phi_{ij}^H = \sum_k \sum_l \Delta^{ij} \varepsilon_{kl}^2, \quad (2.127)$$

amb

$$\Delta^{ij} \varepsilon_{kl}^2 = \varepsilon_{kl}^2|_{x_{ij}=1} - \varepsilon_{kl}^2|_{x_{ij}=0}, \quad (2.128)$$

l'increment respecte d' $(i, j)$  de l'error quadràtic local a  $(k, l)$ . Si mantenim la definició dels increments, quan a l'estat actual  $x_{ij} = 1$ , aleshores  $\varepsilon_{kl}|_{x_{ij}=1} = \varepsilon_{kl}$  i  $\varepsilon_{kl}|_{x_{ij}=0} = \varepsilon_{kl} - \Delta^{ij}\varepsilon_{kl}$ . Però si a l'estat actual  $x_{ij} = 0$ , aleshores  $\varepsilon_{kl}|_{x_{ij}=1} = \varepsilon_{kl} + \Delta^{ij}\varepsilon_{kl}$  i  $\varepsilon_{kl}|_{x_{ij}=0} = \varepsilon_{kl}$ . Així,

$$\Delta^{ij} \varepsilon_{kl}^2 = (1 - 2x_{ij})(\Delta^{ij}\varepsilon_{kl})^2 + 2\varepsilon_{kl}\Delta^{ij}\varepsilon_{kl}, \quad (2.129)$$

i

$$\phi_{ij}^H = (1 - 2x_{ij}) \sum_k \sum_l (\Delta^{ij}\varepsilon_{kl})^2 + 2 \sum_k \sum_l \varepsilon_{kl} \Delta^{ij}\varepsilon_{kl}. \quad (2.130)$$

En termes de xarxes optimitzadores, el camp local equivalent d' $H$  és una expressió que depèn d'uns "estat locals"  $\varepsilon_{kl}$  i uns "pesos"  $\Delta^{ij}\varepsilon_{kl}$ , que es modifiquen a cada actualització (són dinàmics), però que conserven la propietat de simetria ( $\Delta^{ij}\varepsilon_{kl} = \Delta^{kl}\varepsilon_{ij}$ ).

El càlcul de  $\Delta^{ij}\varepsilon_{kl}$  depèn de la posició relativa de  $(k, l)$  i  $(i, j)$ , i dels termes presents en la funció de cost  $F$  original, però no suposa un cost addicional ja que, per tot  $(k, l)$ , l'increment es pot obtenir com a subproducte del càlcul d' $\varepsilon_{ij}$ , com es mostra a continuació. Per a  $(i, j) = (k, l)$ ,  $\Delta^{ij}\varepsilon_{ij} = \Delta^0 F^* - \Delta^1 F^*$ . Per a  $j = l$  (les dues unitats a la mateixa columna), si ens fixem, per exemple, en el cas de les quàdruples, el terme que els correspon a l'expressió de l'error local  $\varepsilon_{ij}$  és

$$\Delta^{ij}N_q = \sum_{k \neq i} \sum_{l \neq j} x_{kj} x_{il} x_{kl}, \quad (2.131)$$

que es pot reescriure com

$$\Delta^{ij}N_q = \sum_{k \neq i} x_{kj} (\Delta^{ij} \Delta^{kj} N_q), \quad (2.132)$$

on

$$\Delta^{ij} \Delta^{kj} N_q = \sum_{l \neq j} x_{il} x_{kl}, \quad (2.133)$$

és el terme corresponent a les quàdruples a l'expressió de l'increment de l'error  $\Delta^{ij}\varepsilon_{kj}$ . Quan les dues unitats estan a la mateixa fila ( $i = k$ ), el terme de les quàdruples és pot rescriure com

$$\Delta^{ij}N_q = \sum_{l \neq j} x_{il} (\Delta^{ij} \Delta^{il} N_q). \quad (2.134)$$

Finalment, si estan en diagonal ( $(i, j) \neq (k, l)$ ), tenim simplement que  $\Delta^{ij} \Delta^{kl} N_q = x_{kj} x_{il}$ .

Al contrari que en les xarxes convencionals, el càlcul incremental de l'error quadràtic no suposa una reducció de la complexitat, ja que cal avaluar tots els errors locals abans d'aplicar l'expressió (2.130). El mateix que caldria per avaluar directament la funció  $H$  (2.124).

El que proposem en aquest apartat, doncs, no és tant un nou model de xarxa, com una alternativa a l'objectiu a optimitzar, que treu partit el coneixement que tenim de l'òptim. La formulació en termes de camp local equivalent, aleshores, ens permetrà aplicar les mateixes estratègies, basades en la cerca local.

## Capítol 3

### Estratègies de cerca

Per a un problema donat, i amb una funció de cost donada, la xarxa optimitzadora del capítol anterior defineix el paisatge de cerca. L'estratègia d'exploració defineix, aleshores, de quina manera recorrem aquest espai per buscar-ne els mínims absoluts, que corresponen a solucions del disseny. En aquest capítol es descriu, des d'una perspectiva algorísmica, el marc general de l'algorisme d'optimització (xarxa + estratègia). I es formulen dins d'aquest marc les tres estratègies que s'avaluen en aquest treball: cerca descendent CD (model de Hopfield), recuita simulada RS (màquina de Boltzmann), i una estratègia pròpia que anomenarem *cerca cooperativa* CC. L'apartat 3.4 descriu i justifica aquest nou model.

#### 3.1 L'algorisme d'optimització. Marc general

Donat un conjunt  $(v, b, u)$  de paràmetres admissible, l'espai de cerca  $\mathcal{A}$  són les  $2^v$  configuracions possibles. Les variables  $x_{ij}$  codifiquen, en cada instant, l'*estat actual*  $A$  de la cerca, i la suma ponderada de les connexions actives avalua l'energia  $E(A)$  d'aquest estat. Els pesos de les connexions codifiquen, com hem vist, la funció de cost escollida, de manera que es compleixi  $E(A) = F(A)$ . En el cas de les xarxes optimitzadores, l'estratègia bàsica d'exploració és la *cerca local*, que consisteix en realitzar únicament transicions entre configuracions *veïnes*. Els *campes locals*  $\phi_{ij}$  avaluen, aleshores, els *increments locals*  $\Delta^i E(A)$  d'energia. Formulats en aquest termes, una estratègia de cerca és un conjunt de criteris per determinar un estat inicial, un ordre d'actualització, una regla de decisió (per acceptar o rebutjar una transició), i una condició de parada. El resultat de la cerca és èxit o fracàs, en funció de si hem trobat o no hem trobat cap solució. Depenent de l'estratègia, alguns d'aquests criteris, com veurem, són paramètrics.

En notació de pseudo-codi, l'esquema general de l'algorisme és el següent:

$V$ , variables de control de la cerca;  
 $A$ , estat actual de la xarxa;  
 $E$ , energia actual;

```

(i, j), unitat actual;
 $x_{ij}$ , estat de la unitat actual;
 $\phi_{ij}$ , camp local de la unitat actual;
 $\Delta_{trans}^{ij}E$ , energia de transició de la unitat actual;

Cerca() {
   $A \leftarrow \text{Estat\_inicial}()$ ;
   $E \leftarrow \text{Energia}(A)$ ;
   $V \leftarrow \text{Valor\_inicial}()$ ;
  repetir {
    iterar (un cop cada unitat) {
      (i, j)  $\leftarrow \text{Següent\_unitat}()$ ;
       $\phi_{ij} \leftarrow \text{Camp\_local}(A, (i, j))$ ;
       $\Delta_{trans}^{ij}E \leftarrow (1 - 2x_{ij})\phi_{ij}$ ;
      si ( $\text{Regla\_de\_decisió}(\Delta_{trans}^{ij}E, V)$ ) {
         $x_{ij} \leftarrow \bar{x}_{ij}$ ;
         $E \leftarrow E + \Delta_{trans}^{ij}E$ ;
      }
       $V \leftarrow \text{Actualitzar\_variables}()$ ;
    }
  } mentre ( $E > \text{Energia\_òptim}()$ )
    i ( $\text{No Condició\_de\_parada}(V)$ );
  si ( $E = \text{Energia\_òptim}()$ ) retorna(èxit);
  sino retorna(fracàs);
}

```

Direm que un *descens* és una execució completa de l'algorisme (acabada en èxit o fracàs). Una *iteració* és l'actualització una vegada de cada una de les unitats de la xarxa. Una *actualització* és l'avaluació del camp local d'una unitat, i l'aplicació de la regla de decisió. I, si la decisió és favorable, anomenarem *transició* a la commutació de la unitat i actualització de l'energia.

En relació a l'esquema anterior, la *xarxa*, que depèn del disseny DB i de la funció  $F$ , es defineixen com

```

Xarxa(DB,  $F$ ) {
  Energia_òptim();
  Energia( $A$ );
  Camp_local( $A, (i, j)$ );
}

```

i les tres operacions  $\text{Energia\_òptim}$ ,  $\text{Energia}$ , i  $\text{Camp\_local}$ , estudiades en el capítol anterior, corresponen, respectivament, a les equacions (2.107, 2.104, 2.108) per a les funcions de cost de tipus  $F$ , i a les equacions (2.125, 2.124, 2.130) per a les funcions de cost de tipus  $H$ .

La definició de l'estratègia és

```

Estratègia( $P_1, P_2, \dots$ ) {
    Ordre d'actualització {
        Estat_inicial();
        Següent_unitat();
    }
    Variables de control {
         $V_1, V_2, \dots$ ;
        Valor_inicial();
        Actualitzar_variables();
    }
    Criteris de relaxació {
        Regla_de_decisió( $\Delta_{trans}^{ij}E, V$ )
        Condició_de_parada( $V$ );
    }
}

```

amb  $P_1, P_2, \dots$  els paràmetres globals de l'estratègia.

En aquest treball, l'Estat\_inicial s'escollirà a l'atzar, i l'operació Següent\_unitat seguirà un ordre d'actualització *seqüencial* prefixat (per files), a fi de restringir tota l'aleatorietat a la primera operació<sup>1</sup>. La resta d'operacions depenen de cada estratègia en particular.

### 3.2 Cerca descendent, CD

La idea bàsica de la cerca descendent, tal com s'ha vist a l'apartat 1.2.1, és acceptar totes les transicions que fan disminuir l'energia de la xarxa, fins que trobem un òptim o l'algorisme *convergeix* en un mínim (després d'una iteració completa, no s'ha acceptat cap transició).

Així, la Condició\_de\_parada és la convergència, i la Regla\_de\_decisió

$$x_{ij} \leftarrow \bar{x}_{ij} \text{ si } \Delta_{trans}^{ij}E < 0, \quad (3.1)$$

depèn únicament del *signe* de  $\Delta_{trans}^{ij}E$ <sup>2</sup>.

La cerca descendent no és paramètrica i la denotarem per CD.

### 3.3 Recuita Simulada, RS

Com hem vist a l'apartat 1.2.2, l'objectiu de la recuita simulada és evitar els mínims locals de la cerca descendent per mitjà de *soroll tèrmic*. Si fem servir la regla

<sup>1</sup>L'alternativa, escollir l'ordre d'actualització a l'atzar, es comenta en el capítol 5.

<sup>2</sup>En realitat, quan  $\Delta_{trans}^{ij}E = 0$  (zones planes), hi ha tres alternatives possibles: no commutar, commutar només si  $x_{ij} = 0$  i commutar només si  $x_{ij} = 1$ . Com veurem al capítol 5, els efectes d'aquesta elecció sobre els resultats de la cerca són significatius.



de Metròpolis, a la *temperatura computacional*  $T$ , la *probabilitat* d'acceptar una transició ve donada per

$$P\{x_{ij} \rightarrow \bar{x}_{ij}\} = \begin{cases} 1, & \Delta_{trans}^{ij} E < 0 \\ e^{-\frac{\Delta_{trans}^{ij} E}{T}}, & \text{altrament.} \end{cases} \quad (3.2)$$

L'eficàcia de la recuita simulada depèn d'una bona *planificació* en el decrement de la temperatura. Donat que l'algorisme és costós, es tracta d'invertir el màxim d'esforç possible al voltant de la *temperatura crítica*  $T_{cri}$ , a la qual es formen les barreres d'energia significatives. El tractament habitual consisteix en partir d'una temperatura inicial  $T_0$  i anar-la decremant, a cada iteració, fins una temperatura final  $T_f$ , segons una llei de decrement del tipus

$$T_n = \tau T_{n-1}, \quad (3.3)$$

amb  $\tau < 1$ . Segons la teoria, després de cada decrement caldria realitzar diverses iteracions per tal de que el sistema retorni a l'equilibri tèrmic. Però, a la pràctica (vegeu, p.ex., [Gutzmann, 87]), els resultats són bons si fem servir una llei de decrement *suau*

$$T_k = \sqrt[b]{\tau} T_{k-1}, \quad (3.4)$$

on el subíndex  $k$  representa *el nombre d'unitats actualitzades*, i  $\tau$  és el coeficient equivalent a una iteració completa. La hipòtesi subjacent és que, prenent decrements petits, amb una granularitat temporal fina, les alteracions de l'equilibri seran poc importants.

Si abans no s'ha trobat cap solució, la *fase estocàstica* de la recuita simulada s'atura després d'un nombre prefixat d'iteracions  $N_t$ . A continuació, però, la cerca continua amb una *fase descendent* (CD), per tal de garantir que arribarem a un mínim.

Així, les variables de control de la recuita simulada són la temperatura  $T$  i el nombre d'iteracions  $N$ , i l'estratègia és paramètrica en  $T_0$ ,  $T_f$  i  $\tau$ . De manera equivalent, els paràmetres que farem servir en aquest treball són el nombre total d'iteracions  $N_t$  donat per

$$N_t = \frac{\ln T_f/T_0}{\ln \tau}, \quad (3.5)$$

La *temperatura central*  $T_c$ , definida com

$$T_c = \sqrt{T_0 T_f}, \quad (3.6)$$

i el rang de temperatures  $\rho$

$$\rho = \frac{T_0}{T_f}, \quad (3.7)$$

que descriu el marge de temperatures explorat. Aquesta formulació dels paràmetres permet de fixar *a priori* l'esforç total invertit en la fase estocàstica de la cerca, i serà útil, en el capítol 5 per a la determinació experimental de la temperatura crítica. Una realització particular d'aquesta estratègia es denotarà per  $RS(T_c, \rho, N_t)$ .

### 3.4 Cerca cooperativa CC, la nova estratègia

La idea central de la recuita simulada és, per mitjà de la reducció progressiva de la temperatura, explorar regions de l'espai de cerca en les que l'energia mitjana és cada vegada més baixa. A cada nova temperatura disminueix l'aleatorietat, i la regió d'exploració es fa més estreta (les barreres de potencial es fan més altes), de manera que la cerca progressa en una seqüència “concèntrica” de “valls” cada cop més profundes. D'aquesta manera, si els òptims es troben efectivament al fons d'una d'aquestes valls, la recuita simulada ens ajuda a trobar-los.

La cerca cooperativa és una estratègia *determinista* que mira de treure partit de la mateixa idea explorant l'espai de cerca en *paral.lel*.

#### 3.4.1 Descripció de l'estratègia

La cerca cooperativa és una estratègia pensada per a sistemes multiprocessador SIMD (molts processadors realitzant simultàniament el mateix algorisme sobre dades diferents). L'objectiu és que, amb un cost de comunicació molt baix, el resultat d' $S$  processadors treballant *cooperativament* sigui millor que el resultat agregat dels mateixos processadors treballant independentment.

L'objectiu és, com abans, fer una exploració progressiva de regions concèntriques d'energia mitjana decreixent. Però enlloc d'explorar l'espai amb una sola instància de la xarxa que, de manera estocàstica, evoluciona cap a estats de baixa energia, el que fem és estimar la mitjana d'una regió mostrejant-la en *paral.lel* amb diverses xarxes (diverses instàncies de la xarxa que exploren estats diferents) que es deixen caure en formació, com una “bola” gran que llisqués cap al fons de la vall. Reduint-ne progressivament el tamany, la “bola” pot endinsar-se cap a regions cada cop més estretes i de menor energia. A aquesta unitat d'exploració l'anomenarem *esquadra*, i a cada una de les  $S$  xarxes que la componen l'anomenarem un *membre* de l'esquadra, i s'executarà en un processador diferent.

L'energia de l'esquadra  $\mathcal{E}$  es defineix com la suma de les energies dels seus membres

$$\mathcal{E} = \sum_{p=1}^S E^p, \quad (3.8)$$

amb  $E^p$  l'energia del membre  $p$ . A cada actualització, seleccionem una unitat  $(i, j)$ , la mateixa en cada un dels membres, i realitzem una transició en cada un d'ells si l'energia de l'esquadra disminueix. Amb l'energia de transició de l'esquadra respecte a la unitat  $(i, j)$  definida com

$$\Delta_{trans}^{ij} \mathcal{E} = \sum_{p=1}^S \Delta_{trans}^{ij} E^p, \quad (3.9)$$

la transició és acceptada d'acord amb la regla de decisió

$$x_{ij}^p \leftarrow \bar{x}_{ij}^p \quad \text{sii} \quad \Delta_{trans}^{ij} \mathcal{E} < 0 \quad \text{per a } p = 1, \dots, S. \quad (3.10)$$

L'esquadra com a tal realitza, doncs, cerca descendent i, a cada transició, malgrat que l'energia individual d'algun dels membres pot augmentar, l'energia global disminueix. Donat que tots els membres commuten la mateixa unitat (tots es mouen en la mateixa direcció), la topologia de l'esquadra es manté constant.

El tamany de la regió d'exploració es regula per mitjà del radi  $R$  de l'esquadra, definit com la màxima distància de Hamming entre el centre (un punt de l'espai de cerca) i qualsevol dels seus membres. De manera similar a la recuita simulada, l'estratègia de la cerca cooperativa es basa en una reducció progressiva del radi de manera que, a cada decrement, l'esquadra s'ha de *contraure* i tots els seus membres s'aproximen cap al centre. L'eficàcia de la cerca, aleshores, depèn d'una bona elecció dels radis inicial  $R_0$  i final  $R_f$ , i d'una bona *planificació* en el mecanisme de decrement. Si no s'ha trobat abans cap solució, la *fase cooperativa* de la cerca s'atura al cap d' $N_t$  iteracions, es desfà l'esquadra, i la cerca continua en *fase descendent* (CD), fins que cada membre convergeixi, independentment, en un mínim. Direm que la *cerca de l'esquadra acaba amb èxit si al menys un dels seus membres troba un òptim*.

L'elecció del tamany  $S$  de l'esquadra també serà rellevant. En aquest treball hem optat per mantenir-lo fix al llarg del procés de cerca, de manera que, a mesura que el radi decreix, la densitat de mostreig de la regió explorada augmenta.

L'elecció de la topologia de l'esquadra i el mecanisme de contracció, quan es redueix el radi, ofereixen diverses possibilitats. Pel que fa a la topologia, la primera alternativa que es proposa consisteix en inicialitzar tots els membres a un mateix estat inicial escollit a l'atzar (el centre de l'esquadra), seleccionar  $R$  unitats a l'atzar, les *mateixes* unitats en cada membre, i assignar-los un valor a l'atzar de manera independent en cada un d'ells. D'aquesta manera tots els membres comparteixen el valor en  $vb - R$  unitats i, en les  $R$  unitats restants (que anomenarem unitats *variables*), mostregen uniformement un subespai de cardinal  $2^R$ . A aquesta topologia l'anomenarem *focalitzada*, perquè concentra tots els membres en un mateix subespai, i la denotarem per CCf.

La segona alternativa, que anomenarem topologia *dispersa* CCd, consisteix en, fixat el centre de l'esquadra, escollir  $R$  unitats *diferents* en cada membre ( $R$  unitats escollides a l'atzar de manera independent), i modificar, com abans, els seus valors a l'atzar. Cada membre té, aleshores, les pròpies unitats variables, i la resta coincideixen amb el centre. Però no existeix un conjunt predefinit d'unitats comuns a tots els membres. En contraposició a la topologia focalitzada, i a l'entorn del punt central, els membres d'una esquadra dispersa es distribueixen uniformement en *qualsevol* de les  $vb$  dimensions, mostrejant uniformement els entorns del punt central, fins a una distància  $R$ .

El *diametre* de l'esquadra, definit com la màxima distància permesa entre dos membres qualsevols, val  $R$  en el cas de les focalitzades i  $2R$  en el cas de les disperses. Però la *distància mitjana efectiva* és  $R/2$  en el primer cas, i tendeix asimptòticament a  $R$ , en el segon cas, quan  $R$  es fa petit en relació a  $vb$ . La *densitat de mostreig* mesura la relació entre el nombre  $S$  de membres del cluster i el tamany de la regió mostrejada, i determina el *radi mínim significatiu*  $R^{\min}$ , per sota del qual

els diferents membres esdevenen redundants. En el cas focalitzat, amb  $S = 2^m$ ,

$$R_f^{min} = m, \quad (3.11)$$

i en el cas dispers, amb  $S < vb$  (la situació més raonable),

$$R_d^{min} = 1. \quad (3.12)$$

La reducció del radi, doncs, no cal que vagi més enllà d' $R^{min}$ .

La *contracció* de l'esquadra, a cada decrement del radi, implica eliminar una unitat variable a cada membre i assignar-li un valor comú. El mecanisme consisteix en escollir a l'atzar una de les  $R$  unitats variables, i decidir a quin valor s'ha de fixar. Per a les esquadres focalitzades, la unitat escollida és la mateixa a tots els membres, i el valor es pot fixar de dues maneres. Si la unitat seleccionada és  $(i, j)$ , la primera alternativa, que anomenarem *contracció descendent*, consisteix en seleccionar el valor (1 o 0) que minimitza l'energia de l'esquadra, segons

$$x_{ij}^p \leftarrow \begin{cases} 1 & \text{si } \Delta^{ij}\mathcal{E} < 0 \\ 0 & \text{altrament} \end{cases} \quad \text{per a } p = 1, \dots, S, \quad (3.13)$$

amb

$$\Delta^{ij}\mathcal{E} = \mathcal{E}|_{x_{ij}^p=1} - \mathcal{E}|_{x_{ij}^p=0} = \sum_{p=1}^S \Delta^{ij}E^p = \sum_{p=1}^S \phi_{ij}^p, \quad (3.14)$$

l'*increment local* d'energia de l'esquadra. El propi mecanisme de contracció, aleshores, contribueix a la cerca descendent. Aquesta variant es denotarà per CCfd.

La segona alternativa, que anomenarem *contracció central* CCfc, consisteix en pendre el membre central  $q$  com a referència, i copiar el seu valor als altres membres. Es a dir,

$$x_{ij}^p \leftarrow x_{ij}^q, \quad \text{per a } p = 1, \dots, S. \quad (3.15)$$

En ambdós casos, l'energia de l'esquadra s'ha d'actualitzar en funció de a quins membres s'ha produït realment una transició. Si denotem per  $y$  el nou valor assignat,

$$\mathcal{E} \leftarrow \mathcal{E} + \sum_{p=1}^S (x_{ij}^p \oplus y) \Delta_{trans}^{ij} E^p, \quad (3.16)$$

amb  $(x_{ij}^p \oplus y)$  l'o-exclusiva entre cada un dels valors vells i el nou valor comú. Així, en el cas de la contracció central, la reducció del radi pot implicar un increment de l'energia de l'esquadra.

Per a les esquadres disperses, donat que les unitats variables són diferents a cada membre, la unitat seleccionada per a fer la contracció s'ha d'escollir independentment a cada membre. Sigui  $(i_p, j_p)$  la unitat seleccionada al membre  $p$ . L'única alternativa factible és ara la contracció central. Així, si  $q$  és el membre de referència

$$x_{i_p j_p}^p \leftarrow x_{i_p j_p}^q \quad \text{per a } p = 1, \dots, S, \quad (3.17)$$

i actualitzem l'energia segons

$$\mathcal{E} \leftarrow \mathcal{E} + \sum_{p=1}^S (x_{ipjp}^p \oplus x_{ipjp}^q) \Delta_{trans}^{ipjp} E^p. \quad (3.18)$$

Per a la planificació del decrement del radi, hem optat per una llei lineal amb granularitat temporal fina<sup>3</sup>. Donat que el radi  $R$  és enter, utilitzarem una variable de control auxiliar  $r \in \mathcal{R}$ , i actualitzarem el radi en funció de la seva part entera. Així,

$$r_k = r_{k-1} - \frac{\Delta r}{vb}, \quad (3.19)$$

amb  $k$  el nombre d'unitats actualitzades, i  $\Delta r$  el decrement equivalent a una iteració completa, i prendrem

$$R_k = \lfloor r_k \rfloor. \quad (3.20)$$

Despres de cada actualització, l'esquadra s'haurà de contraure  $R_{k-1} - R_k$  vegades, que pot ser cap o diverses. Com abans, donats el radi inicial i final, el nombre total d'iteracions de la fase cooperativa és

$$N_t = \frac{R_0 - R_f}{\Delta r}. \quad (3.21)$$

### 3.4.2 Formalització en el marc global

En el cas de la cerca cooperativa, l'algorisme descrit a l'apartat 3.1 es replica en paral·lel  $S$  vegades, i l'única comunicació necessària es produeix en els càlculs relatius a l'energia de l'esquadra. Es tracta, però, d'una comunicació senzilla que consisteix simplement en acumular els termes corresponents de cada un dels  $S$  membres.

La cerca cooperativa és paramètrica, i la denotarem per  $CC(S, R_0, R_f, N_t)$ , amb les tres variants descrites CCfd, CCfc, CCd. Les variables de control de la cerca són, en aquest cas, el paràmetre  $r$  de planificació del radi, el radi  $R$  en sí i el nombre d'iteracions actual  $N$ .

Sense refinar el formalisme, l'energia  $E$  de l'algorisme general representa, en aquest cas, l'energia  $\mathcal{E}$  de l'esquadra, i es calcula segons l'equació (3.8). I el mateix succeix amb l'energia  $\Delta_{trans}^{ij} \mathcal{E}$  de transició (eq. 3.9). Però l'estat  $A$  de la xarxa, la unitat actual  $x_{ij}$  i el camp local  $\phi_{ij}$ , són variables paral·leles amb instàncies independents a cada membre. L'estat inicial s'escull d'acord amb la topologia, però la unitat actual  $(i, j)$  és comú, i la contracció del radi es realitza a cada membre en el moment d'actualitzar variables, segons les expressions de l'apartat anterior. La regla de decisió és descendent en l'energia de l'esquadra i, si no s'ha trobat cap solució en  $N_t$  iteracions, s'acaba amb cerca local.

La cerca cooperativa, amb les seves variants i paràmetres, s'avalua experimentalment al capítol 5.

---

<sup>3</sup>Una alternativa, pràcticament molt costosa, es deixar que l'esquadra convergeixi abans de cada decrement del radi.

## Capítol 4

# Disseny dels experiments per a l'avaluació dels algorismes

Donats un problema i un algorisme (una funció de cost i una estratègia), els resultats elementals són l'èxit (o fracàs) de la cerca, i el cost d'execució de l'algorisme. A l'hora de fer comparacions, la *frequència mitjana d'èxit*, que medeix l'*eficàcia* de la cerca, té en compte únicament el primer d'aquests resultats. De manera que, per incloure'ls tots dos, cal fer una estimació del *cost esperat fins la primera solució*, que en medeix l'*eficiència*. L'*espai d'experimentació* ve definit per tres factors: el repertori de problemes explorat, el repertori de funcions i el repertori d'estratègies de cerca. L'*objectiu* de l'experimentació és, aleshores, determinar quines funcions i estratègies són més aptes per a la generació d'un disseny *qualsevol*.

En aquest capítol es defineixen les mesures experimentals i les seves unitats, i es caracteritzen els seus estadístics en termes del resultat i cost d'un descens. Es descriu l'espai d'experimentació, amb els corresponents factors i subfactors, i es planifiquen els experiments d'avaluació dels algorismes en tres etapes: entrenament, comparacions i aplicació a problemes de tamany creixent.

### 4.1 Mesures experimentals

En aquest apartat es defineix la *iteració* com a unitat de cost per comparar els resultats dels algorismes de cerca, i el cost esperat fins la primera solució com a *resultat* d'un cas experimental particular (un disseny, una funció i una estratègia). Finalment, es proposen mesures per medir-lo, adequades a l'entorn experimental que fem servir.

#### 4.1.1 Cost de relaxació

El *cost* es defineix en aquest treball com el nombre d'*iteracions* realitzades mitjançant l'algorisme de l'apartat 3.1. Aquesta mesura medeix el *cost de relaxació*, que és l'esforç invertit per l'*estratègia* en explorar el *païsatge*, per contraposició al

*temps d'execució*, que depèn de la implementació, i a la *complexitat computacional* de la iteració en si, que és la mateixa per a les tres estratègies. Mentre que el temps d'execució s'ha de medir, i la complexitat computacional es pot analitzar *a priori*, el cost de relaxació depèn de l'habilitat de l'algorisme per trobar una solució. En el cas de la cerca descendent, si no s'ha trobat abans un òptim, el cost de relaxació d'un *descens* depèn del temps de convergència, i en el cas de la recuita simulada i la cerca cooperativa, si bé fixem el nombre total  $N_t$  d'iteracions de les fases estocàstica i cooperativa, respectivament, el valor escollit d' $N_t$  modifica la probabilitat d'èxit. Prescindint, doncs, del temps d'execució (sempre podem fer optimització del codi), i amb la complexitat computacional que s'analitza a continuació, la magnitud que utilitzarem per fer comparacions és el cost de relaxació.

Pel que fa a la complexitat, les tres estratègies de cerca són equivalents. L'operació que domina en l'apartat d'inicialitzacions és el càlcul de l'energia inicial, que depèn de la funció de cost. Per a les funcions de tipus  $F$ , el terme més costós és el corresponent a les quàdruples

$$N_q = \sum_i \sum_{j>i} \sum_k \sum_{l>k} x_{ij} x_{kj} x_{il} x_{kl},$$

present a totes les funcions, que fixa la complexitat en  $\mathcal{O}(v^2 b^2)$ . A cada actualització, el terme principal és l'avaluació del camp local, dominat també per les quàdruples

$$\Delta^{ij} N_q = \sum_{k \neq i} \sum_{l \neq j} x_{kj} x_{il} x_{kl},$$

amb complexitat  $\mathcal{O}(vb)$ . De manera que una iteració completa és  $\mathcal{O}(v^2 b^2)$ , igual que la inicialització.

Per a les funcions de tipus  $H$ , tant el càlcul de l'error quadràtic inicial com el càlcul del camp local equivalent requereixen l'avaluació d' $v \times b$  camps locals, de manera que la seva complexitat és  $\mathcal{O}(v^2 b^2)$ , de manera que la complexitat de la iteració és  $\mathcal{O}(v^3 b^3)$ , un ordre superior al cas de les funcions de tipus  $F$ .

El temps d'execució sí que varia d'una estratègia a l'altra. En la recuita simulada, per exemple, la regla de decisió inclou el càlcul d'una exponencial, i en la cerca cooperativa s'ha d'inicialitzar l'esquadra, i contraure-la a cada decrement del radi. Però aquestes operacions no afecten a la complexitat, ja que no depenen del tamany del problema.

#### 4.1.2 Caracterització dels resultats d'un cas experimental. Nombre esperat d'intents i cost esperat fins la primera solució

Donat que podem identificar fàcilment quan una configuració correspon a un disseny en blocs, el problema de la seva generació algorísmica pertany a la categoria de *problemes de decisió* en els quals el resultat és pot etiquetar d'*èxit* o *fracàs*. Donat que

no pretenem discernir isomorfismes, l'èxit en el nostre cas consisteix en trobar una solució *qualsevol*. Per als algorismes que fem servir, aquest resultat no és determinista, ja que introduïm aleatorietat, com a mínim, en l'elecció de l'estat inicial. En aquestes circumstàncies, donat un cas experimental concret (un disseny, una funció i una estratègia), l'*experiment elemental* és un *descens* de l'algorisme, amb resultat  $x \in \Omega = \{0, 1\} = \{\text{fracàs}, \text{èxit}\}$ , i cost de relaxació  $c$ , en iteracions. La primera,  $x$ , és una variable de Bernouilly, amb probabilitat d'èxit  $p$  i probabilitat de fracàs  $q = 1 - p$ , i la segona,  $c$ , és una variable aleatòria de distribució desconeguda, amb valor esperat  $E(c)$  i varianza  $Var(c)$ .

Si l'elecció de l'estat inicial és aleatòria, les diferents *rèpliques* de l'experiment elemental són *independents* entre si. Definim un *experiment seqüencial* com la repetició de l'experiment elemental tantes vegades com calgui, fins a trobar la primera solució. En aquest experiment,

$$y \equiv \{ \text{Nombre d'intents o descensos fins al primer èxit} \}, \quad (4.1)$$

és una variable aleatòria geomètrica amb paràmetre  $p$  i

$$E(y) = 1/p \quad (4.2)$$

$$Var(y) = q/p^2. \quad (4.3)$$

El valor esperat d' $y$ , o la seva inversa  $p$ , mesuren l'*eficàcia* de la cerca.

Incorporant també el cost, definim

$$z \equiv \{ \text{Cost de relaxació fins al primer èxit} \}, \quad (4.4)$$

que, amb  $c_j$  el cost del descens  $j$ -èssim, es pot expressar com

$$z = \sum_{j=1}^y c_j. \quad (4.5)$$

Aquesta variable permet aleshores mesurar l'*eficiència* de la cerca, i la definim com a *resultat* del cas experimental.

Com mostren els resultats del capítol 5, el cost  $c$  d'un descens depèn del resultat corresponent  $x$ . De manera que, si definim  $b \equiv c|_{x=1}$  com el cost dels descensos "bons" i  $d \equiv c|_{x=0}$  com el cost dels descensos "dolents", aleshores, en l'expressió anterior, tot i que les  $c_j$  són independents entre sí, *no* són independents d' $y$  (només l'últim descens és "bo"). Amb aquestes consideracions, es verifica que

$$E(z) = (E(y) - 1)E(d) + E(b) \quad (4.6)$$

$$Var(z) = Var(y)E^2(d) + (E(y) - 1)Var(d) + Var(b). \quad (4.7)$$

A continuació es descriu el context experimental i les mesures que fem servir per fer una *estimació* d'aquests paràmetres.



### 4.1.3 Estimació dels resultats en un entorn multiprocessador. L'experiment paral.lel

Tots els experiments que es descriuen en aquest treball s'han realitzat en una Connection Machine CM-200 [TMC] amb  $N = 2048$  processadors seqüencials d'un bit, que s'adapten perfectament a les variables d'estat  $x_{ij}$  binàries de la xarxa<sup>1</sup>. Les simulacions s'han organitzat de manera que cada processador realitza un descens de l'algorisme de cerca. Donat que els  $N$  descensos s'executen en paral.lel, tot i que podem disposar del resultat i cost de cada un d'ells, no podem determinar quins intents han permès obtenir quins èxits. En altres paraules, les variables aleatòries  $y$  i  $z$  no són directament observables.

Definim un *experiment paral.lel* com la realització simultània d' $N$  descensos independents. En aquestes circumstàncies, la variable aleatòria,

$$\mathbf{X} \equiv \{ \text{Nombre d'èxits en } N \text{ descensos} \}, \quad (4.8)$$

que es pot expressar com

$$\mathbf{X} = \sum_{i=1}^N \mathbf{x}_i, \quad (4.9)$$

segueix una distribució binomial amb

$$E(\mathbf{X}) = Np \quad (4.10)$$

$$Var(\mathbf{X}) = Npq. \quad (4.11)$$

El millor estimador de la probabilitat d'èxit  $p$  és aleshores

$$\hat{p} = \frac{\mathbf{X}}{N}, \quad (4.12)$$

amb valor esperat  $E(\hat{p}) = p$  i variances  $Var(\hat{p}) = pq/N$ . Els estadístics del nombre de descensos  $y$  es poden estimar, aleshores, en base a l'estimació dels seus paràmetres

$$\hat{E}(y) = 1/\hat{p} \quad (4.13)$$

$$\widehat{Var}(y) = \hat{q}/\hat{p}^2. \quad (4.14)$$

Similarment, els estadístics de  $z$  es poden estimar aleshores per

$$\hat{E}(z) = \left(\frac{1}{\hat{p}} - 1\right)(\bar{d}) + (\bar{b}) \quad (4.15)$$

$$\widehat{Var}(z) = \frac{\hat{q}}{\hat{p}^2}(\bar{d})^2 + \left(\frac{1}{\hat{p}} - 1\right)(S_d^2) + (S_b^2), \quad (4.16)$$

amb  $\bar{b}$  i  $S_b^2$  la mitjana i variances mostrals dels  $X$  descensos bons, i  $\bar{d}$  i  $S_d^2$  la mitjana i variances mostrals dels  $N - X$  descensos acabats en fracàs.

<sup>1</sup>Aquest és un dels motius pels quals no vem escollir la màquina de Boltzmann determinista, que utilitza variables reals, en el nostre repertori d'estratègies.

Amb  $C = \sum_{i=1}^N c_i$ , el cost *total* de l'experiment paral·lel, i amb  $X$ , com abans, el nombre total d'èxits, es verifica que  $\hat{E}(z) = \frac{C}{X}$ . Així, la variable aleatòria

$$w = \frac{C}{X} \quad (4.17)$$

verifica que, per a  $N \rightarrow \infty$ ,

$$E(w) \rightarrow E(z), \quad (4.18)$$

de manera que és un *estimador* asimptòticament consistent d' $E(z)$ . Donat que  $E(z)$  és la nostra mesura d'eficiència, definim  $w$  com la variable *resposta* de l'experiment paral·lel. Les comparacions, aleshores, es realitzaran en termes d' $w$ , i la seva variança  $Var(w)$ , que determina la *precisió*, s'avaluarà experimentalment.

Un cas patològic de l'experiment paral·lel es produeix quan *cap* dels  $N = 2048$  processadors troba un òptim. En aquesta circumstància, que denominarem *fracàs complet*, no podrem fer comparacions quantitatives.

#### 4.1.4 El cas de la cerca cooperativa. Execució d' $S$ descensos en esquadra

En el cas de la cerca cooperativa diem que la cerca acaba en èxit quan *algun* dels  $S$  membres de l'esquadra troba un òptim. De manera que l'experiment elemental és ara un descens de l'esquadra completa. Es a dir, els  $S$  descensos *no independents* dels seus membres.

Els resultats d'aquest nou experiment elemental són  $x_{esq}$  i  $c_{esq}$ , amb  $x_{esq}$  èxit o fracàs *de l'esquadra* i  $c_{esq} = \sum_{p=1}^S c^p$ , el cost total del descens conjunt. A partir d'aquesta definició, paràmetres i estadístiques es deriven com abans i, per a  $M$  esquadres executant en paral·lel, els estimadors d'eficàcia i eficiència es defineixen com

$$\hat{p} = \frac{X_{esq}}{M} \quad (4.19)$$

$$w = \frac{C}{X_{esq}}, \quad (4.20)$$

amb  $X_{esq}$  el nombre d'esquadres exitoses, i  $C = \sum_{e=1}^M \sum_{p=1}^S c^{e,p}$  el cost total.

Direm que la cerca cooperativa és efectiva si els resultats de la *cooperació* entre els  $S$  membres de l'esquadra superen els resultats d' $S$  descensos independents. El cost esperat fins la primera solució és aleshores, efectivament, el paràmetre adequat per fer les comparacions, ja que té en compte al mateix temps els recursos invertits en nombre de descensos, i els recursos invertits en nombre de processadors.

Per tal d'igualar les mostres, els experiments descrits al proper capítol s'han realitzat amb  $M = 2048$  esquadres.

## 4.2 Factors experimentals i planificació dels experiments

Com hem dit a l'apartat 1.3, l'objectiu principal d'aquest treball és l'avaluació dels *algorismes* proposats (funcions més estratègies), aplicats al *problema* de la generació de dissenys en blocs. En termes d'anàlisi experimental l'*espai d'experimentació* ve definit pels nivells dels tres *factors experimentals* principals: problema, funció i estratègia. Cada conjunt admissible de paràmetres d'un disseny constitueix un nivell del factor problema, de manera que tenim infinites instàncies del problema. Les funcions s'organitzen en subfactors, d'acord amb la seva estructura, tal com s'ha vist a l'apartat 2.4.3 i, per a cada estructura, els nivells venen donats pels coeficients de composició (novament, infinites opcions). Finalment, les diferents estratègies del capítol 3 també es formulen com a subfactors (CD, RS, CC), amb les tres variants de la cerca cooperativa definides com a subfactors addicionals de l'estratègia CC. Per a les estratègies paramètriques (totes excepte CD), el valor (real) dels paràmetres defineix els (infinites) nivells. Tenint en compte tot això, una anàlisi exhaustiva és, òbviament, intractable, i l'experimentació es planifica en tres *etapes*: una etapa d'*entrenament*, per a seleccionar funcions i ajustar paràmetres, una etapa de *comparacions*, per avaluar les prestacions de cada factor, i una *tercera* etapa on s'aplica el millor algorisme trobat a problemes de tamany creixent.

Pel que fa als dissenys, els 129 casos més petits de [Mathon & Rosa, 90] es llisten a la taula 4.1, i corresponen a tots els conjunts possibles de paràmetres  $(v, b, u)$  admissibles amb  $vb \leq 1000$  i  $3b \leq u \leq vb/2$  (els dissenys amb un nombre d'uns més petit són trivials, mentre que els casos amb  $u > vb/2$  ja estan representats pel seu corresponent complementari). Els problemes estan ordenats per tamany, amb  $d_i$  el número de problema,  $m_j$  la numeració original,  $(v, b, u)$  els paràmetres del disseny,  $[r, k, \lambda]$  els descriptors restants,  $vb$  el tamany del problema, i  $N_s$  les cotes conegudes al nombre de solucions no isomorfes que existeixen (apartat 1.1). El problema d94, amb  $vb = 726$ , és el cas més petit de problema obert (l'existència del disseny no està establerta). El fet és rellevant ja que demostra que, avui per avui, els problemes d'aquest tamany encara no es poden analitzar exhaustivament.

L'etapa de comparacions es realitza sobre els 25 primers problemes, els més petits, que constitueixen el que anomenarem el conjunt de *prova*. D'entre ells, 7 casos escollits a l'atzar  $\{d2, d4, d8, d10, d15, d18, d21\}$  formen el conjunt d'*entrenament*. Els 104 problemes restants s'utilitzen a la tercera etapa.

L'*etapa d'entrenament* s'organitza de la següent manera. Primer es comparen les funcions de cost entre si per seleccionar les millors. A fi de restringir la casuística s'utilitza únicament CD com a estratègia de *referència*, i es fan els experiments amb el conjunt de problemes d'entrenament. De les 9 possibles estructures, 2 tenen la composició fixada. Per a cada una de les altres 7, es proven diferents combinacions de coeficients, i s'escull en cada cas la combinació que millors resultats dona. Finalment, la comparació entre les 9 funcions obtingudes permet seleccionar les funcions que anomenarem *finalistes*.

A continuació s'ajusten els paràmetres de les estratègies amb els problemes del conjunt d'entrenament, prenent  $F_{uq}$ , la més senzilla com a funció de *referència*. CD no té paràmetres, RS depèn de  $(T_c, \rho, N_t)$ , i CC depèn de  $(S, R_0, R_f, N_t)$ . Les variants de CC s'ajusten per separat, i es comparen entre si al final per escollir la millor. Per a cada estratègia, es realitza una optimització experimental sobre cada un dels problemes, que condueix a un conjunt de paràmetres *òptim* per a aquell problema. Els paràmetres resultants, aleshores, es *generalitzen* respecte a funcions i problemes, per obtenir el conjunt de paràmetres *estàndard* de l'estratègia. Aquest és el valor dels paràmetres que farem servir en les següents etapes.

L'*etapa de comparacions* es realitza aleshores sobre un espai experimental reduït: el conjunt de problemes de prova (que inclou, com hem dit, el conjunt d'entrenament), les funcions de cost finalistes, i les tres estratègies amb paràmetres estàndard. Així podem provar totes les combinacions. Aleshores es determina la millor funció de cost per a cada estratègia, es comparen les estratègies entre si (tant en termes d'eficàcia com en termes d'eficiència), i s'analitza la dificultat dels problemes en funció del seu tamany.

Finalment, a la *tercera etapa*, s'aplica el millor algorisme trobat (la millor combinació funció-estratègia) als 104 problemes restants, per ordre de tamany creixent.

Degut al procediment experimental que fem servir, les decisions preses a l'etapa d'entrenament no estan completament contrastades, i pot haver-hi alguna combinació més bona. Els resultats de les comparacions, aleshores, per molt clars que puguin semblar, no es poden extrapolar a la lleugera.

$d_i$	$m_i$	$v$	$b$	$u$	$r$	$k$	$\lambda$	$vb$	$N_s$
d1	m1	7	7	21	3	3	1	49	1
d2	m4	6	10	30	5	3	2	60	1
d3	m9	7	14	42	6	3	2	98	4
d4	m2	9	12	36	4	3	1	108	1
d5	m15	8	14	56	7	4	3	112	4
d6	m43	6	20	60	10	3	4	120	4
d7	m7	11	11	55	5	5	2	121	1
d8	m31	7	21	63	9	3	3	147	10
d9	m10	10	15	60	6	4	2	150	3
d10	m24	9	18	72	8	4	3	162	11
d11	m3	13	13	52	4	4	1	169	1
d12	m33	10	18	90	9	5	4	180	21
d13	m118	6	30	90	15	3	6	180	6
d14	m67	7	28	84	12	3	4	196	35
d15	m21	9	24	72	8	3	2	216	36
d16	m101	8	28	112	14	4	6	224	2224
d17	m20	15	15	105	7	7	3	225	5
d18	m236	6	40	120	20	3	8	240	13
d19	m47	11	22	110	10	5	4	242	4393
d20	m117	7	35	105	15	3	5	245	108
d21	m13	16	16	96	6	6	2	256	3
d22	m58	12	22	132	11	6	5	264	601
d23	m191	7	42	126	18	3	6	294	417
d24	m71	10	30	120	12	4	4	300	> 998
d25	m30	10	30	90	9	3	2	300	960
d26	m409	6	50	150	25	3	10	300	19
d27	m16	15	21	105	7	5	2	315	0
d28	m5	16	20	80	5	4	1	320	1
d29	m66	9	36	108	12	3	3	324	22521
d30	m150	9	36	144	16	4	6	324	$\geq 12$
d31	m278	8	42	168	21	4	9	336	$\geq 943$
d32	m23	13	26	104	8	4	2	338	2407
d33	m77	13	26	156	12	6	5	338	$\geq 1017$
d34	m8	13	26	78	6	3	1	338	2
d35	m276	7	49	147	21	3	7	343	$\geq 9$
d36	m195	10	36	180	18	5	8	360	$\geq 1000$
d37	m596	6	60	180	30	3	12	360	34
d38	m41	19	19	171	9	9	4	361	6
d39	m125	11	33	165	15	5	6	363	$\geq 127$
d40	m89	14	26	182	13	7	6	364	$\geq 79$
d41	m35	16	24	144	9	6	3	384	$\geq 1512$
d42	m357	7	56	168	24	3	8	392	$\geq 35$
d43	m56	12	33	132	11	4	3	396	$\geq 10^3$

Taula 4.1: Paràmetres  $(v, b, u)$  admissibles per a  $vb \leq 1000$  i  $3b \leq u \leq vb/2$ . Amb  $d_i$  numeració pròpia,  $m_i$  numeració original,  $(v, b, u)$  paràmetres del disseny,  $[r, k, \lambda]$  descriptors,  $vb$  tamany del problema i  $N_s$  nombre de solucions no isomorfes conegudes.  $N_s = 0$  vol dir que el disseny no existeix, i  $N_s = ?$  indica un problema obert. (De [Mathon & Rosa, 90]).

$d_i$	$m_i$	$v$	$b$	$u$	$r$	$k$	$\lambda$	$vb$	$N_s$
d44	m816	6	70	210	35	3	14	420	48
d45	m145	9	48	144	16	3	4	432	$\geq 330$
d46	m6	21	21	105	5	5	1	441	1
d47	m477	7	63	189	27	3	9	441	$\geq 10$
d48	m275	8	56	168	21	3	6	448	$\geq 101$
d49	m524	8	56	224	28	4	12	448	$\geq 2224$
d50	m193	10	45	180	18	4	6	450	$\geq 14819$
d51	m109	15	30	210	14	7	6	450	$\geq 11$
d52	m130	16	30	240	15	8	7	480	$\geq 9 \times 10^7$
d53	m1078	6	80	240	40	3	16	480	76
d54	m247	11	44	220	20	5	8	484	$\geq 4394$
d55	m19	22	22	154	7	7	2	484	0
d56	m364	9	54	216	24	4	9	486	$\geq 10^6$
d57	m595	7	70	210	30	3	10	490	$\geq 108$
d58	m70	13	39	156	12	4	3	507	$\geq 10^3$
d59	m124	13	39	195	15	5	5	507	$\geq 30$
d60	m76	16	32	192	12	6	4	512	$\geq 111$
d61	m14	15	35	105	7	3	1	525	80
d62	m104	15	35	210	14	6	5	525	$\geq 117$
d63	m55	12	44	132	11	3	2	528	$\geq 10^6$
d64	m319	12	44	264	22	6	10	528	$\geq 602$
d65	m63	23	23	253	11	11	5	529	1103
d66	m735	7	77	231	33	3	11	539	$\geq 107$
d67	m480	10	54	270	27	5	12	540	$\geq 10^8$
d68	m235	9	60	180	20	3	5	540	$\geq 330$
d69	m819	8	70	280	35	4	15	560	$\geq 2224$
d70	m158	17	34	272	16	8	7	578	$\geq 11$
d71	m25	21	28	168	8	6	2	588	0
d72	m881	7	84	252	36	3	12	588	$\geq 417$
d73	m190	10	60	180	18	3	4	600	$\geq 961$
d74	m363	10	60	240	24	4	8	600	$\geq 14819$
d75	m116	11	55	165	15	3	3	605	$\geq 436800$
d76	m242	11	55	220	20	4	6	605	$\geq 1$
d77	m416	11	55	275	25	5	10	605	$\geq 3337$
d78	m179	18	34	306	17	9	8	612	$\geq 10^3$
d79	m40	25	25	225	9	9	3	625	78
d80	m102	15	42	210	14	5	4	630	$\geq 103$
d81	m49	21	30	210	10	7	3	630	$\geq 414$
d82	m1030	7	91	273	39	3	13	637	$\geq 417$
d83	m44	16	40	160	10	4	2	640	$\geq 986$
d84	m128	16	40	240	15	6	5	640	$\geq 15$
d85	m356	9	72	216	24	3	6	648	$\geq 10^7$
d86	m690	9	72	288	32	4	12	648	$\geq 10^6$

Taula 4.1: (Continuació).

$d_i$	$m_i$	$v$	$b$	$u$	$r$	$k$	$\lambda$	$vb$	$N_s$
d87	m290	15	45	315	21	7	9	675	$\geq 10^8$
d88	m65	13	52	156	12	3	2	676	$\geq 92714$
d89	m149	13	52	208	16	4	4	676	$\geq 2408$
d90	m373	13	52	312	24	6	10	676	$\geq 1018$
d91	m891	10	72	360	36	5	16	720	$\geq 10^8$
d92	m208	19	38	342	18	9	8	722	$\geq 7$
d93	m608	11	66	330	30	5	12	726	$\geq 10^6$
d94	m78	22	33	264	12	8	4	726	?
d95	m451	14	52	364	26	7	12	728	$\geq 80$
d96	m98	27	27	351	13	13	6	729	$\geq 7$
d97	m131	21	35	315	15	9	6	735	$\geq 10^4$
d98	m599	10	75	300	30	4	10	750	$\geq 29638$
d99	m11	25	30	150	6	5	1	750	1
d100	m517	9	84	252	28	3	7	756	$\geq 330$
d101	m224	20	38	380	19	10	9	760	$\geq 10^{16}$
d102	m123	16	48	240	15	5	4	768	$\geq 11$
d103	m200	16	48	288	18	6	6	768	$\geq 10^8$
d104	m316	12	66	264	22	4	6	792	$\geq 10^3$
d105	m743	12	66	396	33	6	15	792	$\geq 602$
d106	m1088	9	90	360	40	4	15	810	$\geq 10^6$
d107	m28	29	29	232	8	8	2	841	0
d108	m241	13	65	260	20	4	5	845	$\geq 10^3$
d109	m826	11	77	385	35	5	14	847	$\geq 10^6$
d110	m683	9	96	288	32	3	8	864	$\geq 10^7$
d111	m46	21	42	210	10	5	2	882	$\geq 10$
d112	m75	21	42	252	12	6	3	882	$\geq 1$
d113	m259	21	42	420	20	10	9	882	$\geq 4$
d114	m284	16	56	336	21	6	7	896	$\geq 1$
d115	m476	10	90	270	27	3	6	900	$\geq 10^{12}$
d116	m888	10	90	360	36	4	12	900	$\geq 10^9$
d117	m537	15	60	420	28	7	12	900	$\geq 10^{18}$
d118	m176	18	51	306	17	6	5	918	$\geq 3$
d119	m293	22	42	462	21	11	10	924	$\geq 2$
d120	m280	15	63	315	21	5	6	945	$\geq 2211$
d121	m119	16	60	240	15	4	3	960	$\geq 6 \times 10^5$
d122	m618	16	60	480	30	8	14	960	$\geq 9 \times 10^7$
d123	m12	31	31	186	6	6	1	961	1
d124	m54	31	31	310	10	10	3	961	$\geq 38$
d125	m143	31	31	465	15	15	7	961	$\geq 1266891$
d126	m1095	11	88	440	40	5	16	968	$\geq 10^7$
d127	m108	22	44	308	14	7	4	968	$\geq 1$
d128	m880	9	108	324	36	3	9	972	$\geq 330$
d129	m160	25	40	400	16	10	6	1000	$\geq 43$

Taula 4.1: (Continuació).

# Capítol 5

## Descripció dels experiments i anàlisi dels resultats

En aquest capítol es descriuen els experiments realitzats per avaluar la dificultat dels problemes, la qualitat de les funcions de cost, i l'eficiència de les estratègies de cerca. En la primera etapa es caracteritzen estadísticament les variables resultat  $z$  i resposta  $w$ , es fa una preselecció de les funcions de cost, i s'ajusten els paràmetres de les diferents estratègies. En la segona etapa es busquen les millors funcions de cost, es comparen i validen les estratègies, i s'analitzen els resultats en funció del tamany del problema. Finalment, en la tercera etapa, s'aplica el millor algorisme trobat a problemes cada vegada més grans.

### 5.1 Primera etapa: Entrenament

#### 5.1.1 Experiments preliminars. Caracterització de $z$ i $w$

L'objectiu d'aquest apartat és fer una anàlisi de la distribució estadística de les variables resultat i resposta, i contrastar-la amb l'estimació dels seus paràmetres donada pels models de l'apartat 4.1. L'anàlisi es realitza per a un sol disseny d15 (relativament gran, però relativament fàcil), per a una sola funció  $F_{uq}$  (la millor de l'apartat 5.3.1), i per a una instància de cada estratègia de cerca (amb els paràmetres estàndard i variants de l'apartat 5.2).

Donat que ambdues variables són funció del cost  $c$  d'un descens, de distribució desconeguda, el primer experiment que presentem caracteritza el comportament d'aquesta variable condicionada a l'èxit o fracàs del resultat del descens.

#### Experiment 1

**Objectius:** Anàlisi de la distribució del cost  $b$  d'un descens acabat en èxit, i del cost  $d$  d'un descens acabat en fracàs.

**Problemes:** d15.



**Funcions:**  $F_{uq}$ .

**Estratègies:** CD,  $RS^{std}$ ,  $CC^{std}$ .

**Rèpliques:** 2048, de l'experiment elemental (un descens). El nombre de "bons" i "dolents" dependrà dels resultats.

**Descripció:** Un descens a cada processador. Per a cada un d'ells anotem el resultat  $x$  (òptim-local) del descens, i el cost corresponent ( $b-d$ ).

La figura 5.1, al final d'aquest capítol, mostra els histogrames dels resultats obtinguts, contrastats en cada cas amb una distribució normal de la mateixa mitjana i variança. La constatació més evident és la diferència de comportament entre les tres estratègies. Per a la cerca descendent, tant la distribució del cost dels òptims com la dels mínims locals s'inclina lleugerament cap a l'esquerra, però ambdues s'allunyen poc de la normal. La diferència en mitjanes és petita però, tenint en compte el tamany de les poblacions, és significativa.

Per a la recuita simulada, les distribucions es comporten de la manera esperada. Si recordem el funcionament de l'algorisme, el nombre *màxim* d'iteracions a  $T \neq 0$  és fixe ( $N_t = 100$ , en aquest cas) i el descens continua fins a la convergència amb cerca descendent. Els òptims, però, s'aturen tan bon punt es detecten. La distribució de  $b$  demostra que pràcticament tots els òptims apareixen en la fase estocàstica del descens. I la distribució que obtenim, desviada clarament cap a l'esquerra, recorda una distribució de Poisson (temps de vida d'una execució fins a caure en un mínim global). Pel que fa als mínims locals, pràcticament tots acaben a la iteració 101 (cal una iteració addicional per verificar que és un mínim), de manera que també s'han format en la fase estocàstica. Amb els paràmetres utilitzats, doncs, la recuita simulada fa pràcticament tota la feina.

En el cas de la cerca cooperativa, finalment, amb un esquema com l'anterior (fase cooperativa + cerca descendent), amb  $N_t = 15$  i  $S = 2$ , el cost màxim de la primera fase és 30. Contràriament al cas de la recuita simulada, els resultats obtinguts il·lustren que pràcticament tots els mínims (globals i locals) es formen a la segona fase. La millora en el rediment (361 òptims respecte als 72 de la CD) s'ha d'atribuir, aleshores, a que la fase cooperativa "millora" els estats inicials. Les distribucions obtingudes, doncs, són qualitativament iguals a les de CD, amb 30 iteracions de més.

El següent experiment compara les distribucions d' $y$  i  $z$ .

## Experiment 2

**Objectius:** Anàlisi de la distribució del nombre de descensos  $y$  i del cost  $z$  fins la primera solució.

**Problemes:** d15.

**Funcions:**  $F_{uq}$ .

**Estratègies:** CD,  $RS^{std}$ ,  $CC^{std}$ .

**Rèpliques:** 2048, de l'experiment seqüencial (fins òptim).

**Descripció:** Per a cada processador, fem tants descensos com calgui fins a trobar un òptim, i anotem, en cada cas, el cost total  $z$ , el nombre de descensos  $y$ , el cost de l'últim descens  $b$ , i el cost acumulat (lineal i quadràtic) de la resta de descensos  $d_j$ .

La figura 5.2 mostra l'histograma de les distribucions obtingudes per a cada una de les estratègies. Per a CD, amb una distribució de costos pràcticament normal, la distribució de  $z$  és qualitativament molt semblant a la d' $y$ . Per a RS, l'eficàcia de la cerca és tan gran, en aquest cas, que la major part dels òptims s'assoleixen en un intent, i mana la distribució de  $b$ . Una observació més acurada mostra una segona corba a partir de la iteració 100. Donat que  $d$  és pràcticament constant i  $b < d$ , la distribució de  $z$  presenta, per a cada valor d' $y$  (i amb l'amplitut corresponent), una rèplica de la distribució de  $b$ . Finalment, el comportament de CC és qualitativament similar al de CD.

Les mitjanes i desviacions mostrals obtingudes és tabulen a continuació.

	$\bar{y}$	$S_y$	$\bar{z}$	$S_z$
CD	27.63	26.61	158.63	151.79
RS	1.11	0.34	42.03	39.30
CC	5.40	5.04	189.40	176.29

El següent experiment analitza la distribució de la variable resposta  $w$  de l'experiment paral.lel, i l'estimació que s'en deriva dels estadístics de  $z$ .

**Experiment 3**

**Objectius:** Anàlisi de la distribució de la resposta  $w$ .

**Problemes:** d15.

**Funcions:**  $F_{uq}$ .

**Estratègies:** CD, RS<sup>std</sup>, CC<sup>std</sup>.

**Rèpliques:** 51, de l'experiment paral.lel.

**Descripció:** A cada rèplica (un descens a cada processador) anotem el nombre total d'èxits  $X$ , i el cost acumulat (lineal i quadràtic) dels descensos acabats en èxit i dels descensos acabats en fracàs.

La figura 5.3 mostra els histogrames obtinguts per a cada una de tres les estratègies. Els resultats obtinguts mostren una lleugera desviació cap a l'esquerra de les distribucions, i una dispersió molt baixa en el cas de RS. Més endavant estudiarem més en detall aquests fenòmens.

Pel que fa a la qualitat dels estimadors, amb les dades de l'experiment paral.lel obtenim 51 valors diferents per a cada un dels estimadors de la mitjana i la variança (expressada en termes de desviació) d' $y$  i  $z$ , basats en les expressions de

l'apartat 4.1.3 (equacions (4.13, 4.14) i (4.15, 4.16), respectivament). La taula següent mostra, en cada cas, la mitjana i la desviació (entre parèntesi) d'aquests 51 valors.

	$\hat{E}(z)$		$\widehat{Dev}(z)$		$\hat{E}(y)$		$\widehat{Dev}(y)$	
CD	28.90	(3.31)	28.40	(3.31)	165.91	(19.08)	162.36	(19.07)
RS	1.10	(0.01)	0.33	(0.01)	41.01	(0.92)	38.04	(1.23)
CC	5.44	(0.24)	4.92	(0.24)	190.82	(8.40)	171.99	(8.44)

Recordem que, per definició,  $\hat{E}(z)$  és el resultat  $w$  de l'experiment paral·lel. Comparades amb la taula anterior, les estimacions són prou bones.

L'anàlisi de la variabilitat de  $w$  és important, ja que és la variable que farem servir per comparar algorismes i problemes entre si. L'anàlisi estadística és, en general, molt sensible a les desviacions de la normalitat en la variable resposta, i la majoria de tests es basen en la hipòtesi de variança constant. Els resultats que es presenten a continuació mostren com, en el nostre cas, els requisits anteriors no es compleixen.

Pel que fa a la distribució, el següent experiment mostra alguns exemples més.

Experiment 4

- Objectius: Anàlisi de la istribució de la resposta  $w$ .
- Problemes: d5,d10,d7,d11.
- Funcions:  $F_{uq}$ .
- Estratègies: CD.
- Rèpliques: 30 (120 per a d10).

La figura 5.4 mostra els histogrames obtinguts. L'anàlisi d'aquesta figura (i la de la figura anterior), no permet treure conclusions definitives. Però sí que es pot observar que, en alguns casos (com a) i b) a la figura 5.4, i c) a la figura 5.3), les desviacions de la normalitat semblen significatives. La distribució de  $w$ , a més, sembla que varia d'un cas a l'altre. Els exemples presentats aquí, escollits arbitràriament en la fase preliminar de l'experimentació, no són representatius de tota la casuística, però il·lustren el tipus de problemes amb que ens podem trobar.

La hipòtesi d'igualtat de variàncies definitivament no es compleix. La figura 5.5a) presenta un gràfic de la desviació mostral respecte a la mitjana, per als 7 problemes del conjunt d'entrenament, amb  $F_{uq}$  i les tres estratègies òptimes (15 rèpliques en cada cas). Les dades corresponen a l'experiment 12, que es descriu més endavant. I la figura 5.5b) presenta el mateix gràfic per als 25 problemes de l'experiment 15, amb paràmetres estàndard (5 rèpliques). Els casos de fracàs complet no es presenten. S'observen dos fenòmens. D'una banda, el comportament és diferent per a les diferents estratègies. I, d'altra banda, tot i créixer clarament amb la mitjana, les desviacions evolucionen de manera irregular. De manera que no podem trobar cap transformació que les iguali.

Recapitulant, doncs, les estimacions d' $E(z)$  i  $Var(z)$  obtingudes per mitjà de l'experiment paral·lel mostren una *exactitud* acceptable (comparades amb la mesura directa dels paràmetres estimats, per mitjà de l'experiment seqüencial). Però la variabilitat de la resposta  $w$ , que determina la *precisió* de la mesura paral·lela, amb una distribució possiblement no normal, una variança variable, i casos de fracàs complet, no permeten l'aplicació de les tècniques estadístiques objectives habituals (test ANOVA, superfície de resposta, etc.).

En els experiments que segueixen, per determinar quina de dues respostes és més bona, els criteris que farem servir són els següents. En primer lloc les comparacions es faran en termes del nombre de fracassos complets (en relació al nombre de rèpliques de l'experiment) i, només en cas d'igualtat, considerarem la resposta  $w$ . Aleshores, en el darrer cas, direm que la diferència entre mitjanes és *significativa* únicament si el seu valor absolut supera la *suma* de les desviacions corresponents. Comparat amb els marges de confiança per a la diferència entre mitjanes en el cas de variables gaussianes, aquest criteri és força precís per a  $n = 15$  rèpliques, i resulta suficient, per als nostres propòsits, per a  $n = 5$ .

### 5.1.2 Preselecció de les funcions de cost

La preselecció de funcions s'ha realitzat en base únicament a l'estratègia CD, i les dues funcions que en resulten *finalistes* s'avaluen a l'apartat 5.2.1 per a la resta d'estratègies de cerca.

Amb el primer experiment busquem una bona composició per a cada una de les nou estructures de les funcions del capítol 2.

#### Experiment 5

**Objectius:** Cerca d'una bona composició per a cada estructura.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}, F_{utq}, F_{uhq}, F_{uvq}, F_{uthq}, F_{utvq}, F_{uhvq}, F_{uthvq}, F_{q\bar{q}}$ .

**Estratègies:** CD.

**Rèpliques:** 1.

**Descripció:** Exceptuant  $F_{uq}$  i  $F_{q\bar{q}}$ , que són fixes, provem diferents valors dels coeficients de composició per a cada estructura, i anotem la resposta  $w$ . Iterativament, amb les dades obtingudes, es busca una nova composició que millori els resultats.

L'elecció dels coeficients s'ha realitzat manualment, en base a l'experiència adquirida, i en la major part dels casos s'han provat únicament 2 o 3 conjunts de valors. Els cas explorat més a fons (unes 30 composicions diferents), correspon a l'estructura  $F_{uthvq}$ , que ens farà servei en el capítol 6. Els resultats detallats no es presenten, però el tipus de problemes que apareixen s'il·lustren a continuació. La taula següent mostra els resultats obtinguts amb 3 composicions diferents d' $F_{uthvq}$ , per als problemes d4 i d8.

$\rho_t$	$\rho_h$	$\rho_v$	$\rho_q$	d4	d8
1	9	3	5	-	56*
1	7	3	9	81.5*	-
1	9	3	7	73.1	24.3

Els guions indiquen fracàs complet, i els asteriscs senyalen els millors resultats obtinguts *en tota l'experimentació*. Com s'observa, l'optimització de coeficients no es pot realitzar cas a cas. I si es vol resoldre el major nombre de problemes de manera "acceptable", cal arribar a un compromís. L'opció escollida, en aquest cas, és la de la tercera fila.

A continuació, fent servir com a criteri la minimització del nombre de fracassos complets, i en base els resultats de l'experiment anterior, es fa una preselecció de les 4 funcions de cost següents:

$$\begin{aligned}
 &F_{uq}(\alpha_u, 0, 0, 0, 1, 0) \\
 &F_{uthq}(\alpha_u, 1, 2, 0, 3, 0) \\
 &F_{uhvq}(\alpha_u, 0, 1, 1, 8, 0) \\
 &F_{uthvq}(\alpha_u, 1, 9, 3, 7, 0).
 \end{aligned} \tag{5.1}$$

Les alternatives eliminades són significativament inferiors. I les funcions escollides, que anomenem *prefinalistes*, s'avaluen amb més precissió a l'experiment següent:

#### Experiment 6

**Objectius:** Comparació entre funcions.

**Problemes:** Conjunt d'entrenament.

**Funcions:** Les 4 funcions prefinalistes.

**Estratègies:** CD.

**Rèpliques:** 15.

Els resultats obtinguts es mostren a la taula de la figura 5.6. Tal com dèiem a l'experiment anterior, el fenomen més remarcable que s'observa és la forta interacció entre funcions i problemes (algunes van millor en alguns problemes que altres). Resulta difícil decidir quina és millor.

La decisió que prenem, en aquest cas, és eliminar les dues funcions del mig superades, respectivament, per les funcions dels extrems de la taula. I queden com a *finalistes*

$$\begin{aligned}
 &F_{uq}(\alpha_u, 0, 0, 0, 1, 0) \\
 &F_{uthvq}(\alpha_u, 1, 9, 3, 7, 0).
 \end{aligned} \tag{5.2}$$

Pel que fa a les funcions de tipus *H*, ens remitim als resultats exposats a [Fontdecaba, 94], que il·lustren que a efectes d'optimització combinatòrica, el seu comportament és molt pobre, fins i tot per als problemes més petits.

### 5.1.3 Optimització experimentals dels paràmetres de la RS

El funcionament de la recuita simulada  $RS(T_c, \rho, N_t)$  depèn de tres paràmetres: la temperatura central d'exploració  $T_c$ , la relació  $\rho$  entre les temperatures inicial i final, i el nombre total d'iteracions  $N_t$  de la fase estocàstica. La hipòtesi de treball és que si el rang de temperatures explorat conté la temperatura crítica  $T_{cri}$ , i el temps invertit és adequat, el cost fins la primera solució ha de ser mínim. El primer experiment, que es descriu a continuació, busca l'òptim d'aquest paràmetres per a cada un dels problemes del conjunt d'entrenament.

#### Experiment 7

**Objectius:** Optimització dels paràmetres de RS.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ .

**Estratègies:**  $RS(T_c, \rho, N_t)$ .

**Rèpliques:** 1.

**Descripció:** De manera independent per a cada problema, optimització seqüencial d' $w$  respecte als paràmetres de l'estratègia (veure text).

Les proves preliminars mostren, efectivament, l'existència d'una temperatura crítica. Fixat el nombre d'iteracions  $N_t$ , si la  $T_c$  escollida és bona, la millor resposta es produeix quan més petit és  $\rho$  (tot l'esforç s'inverteix a la temperatura adequada). De manera que, deixant un marge de tolerància, obtem per fixar  $\rho = 2$  i optimitzar respecte als altres dos paràmetres.

L'estratègia d'optimització que fem servir és la següent: A partir d'un punt de prova  $(T_c^1, N_t^1)$ , escollit a l'atzar, definim un increment per a cada un dels paràmetres,  $(\Delta T_c, \Delta N_t)$ , i fem un experiment factorial amb totes les combinacions  $(T_c^1 \pm \Delta T_c, N_t^1 \pm \Delta N_t)$ . Si la resposta al punt de prova és significativament més bona que a la resta, definim aquest punt com a òptim. Altrament, repetim el procés prenent com a nou punt central l'extrem que millors resultats ha produït.

A la pràctica, els resultats que s'obtenen no mostren un òptim clar. El marge de valors de  $T_c$  que optimitzen la resposta  $w$  és, a vegades, força ampli. Però a banda i banda d'aquest interval els resultats empitjoren. Pel que fa al nombre d'iteracions els efectes són menys importants. En la majoria dels casos, l'augment en la probabilitat d'èxit per a  $N_t$  creixent (refredament lent) compensa el cost addicional invertit, de manera que els resultats no decauen. Si tenim en compte, a més, la variabilitat de les dades, l'òptim no es localitza en un punt, sino en un marge de valors. Per citar un parell d'exemples, en el cas de d10 els resultats són bons per a  $T_c \in [0.64, 0.86]$  i  $N_t > 40$ . I en el cas de d21, s'obtenen bons resultats amb  $T_c \in [0.99, 1.47]$  i  $N_t > 70$ . Un cas excepcional correspon al problema d2, que amb CD ja presenta tan bons resultats que la RS no aconsegueix millorar-los. El millor punt de treball és, aleshores,  $N_t = 1$  (el més pròxim a CD), i és pràcticament insensible a  $T_c$ .

Amb les dades anteriors, escollim manualment el punt *òptim* de treball  $(T_c^*, N_t^*)$  per a cada problema. Per minimitzar el temps d'experimentació, els valors d' $N_t$  escollits són el més petits possible, sense sortir de la regió on els resultats són més bons. Els valors escollits es mostren a la taula de la figura 5.7. I configuren el que denominarem la RS *òptima*

$$RS^{opt} \equiv RS(T_c^*, 2, N_t^*). \quad (5.3)$$

Com s'observa, els valors escollits difereixen notablement d'un disseny a l'altre.

Per definir uns paràmetres estàndard, hem de trobar un mecanisme que generalitzi aquests valors a qualsevol disseny i funció. Pel que fa a  $T_c$ , la regla de decisió de la RS (eq. 3.2) depèn de la *relació* entre els camps locals i la temperatura. Al llarg de la recuita, el moment en que comencen a aparèixer solucions correspon a la temperatura crítica  $T_{cri}$ , i es verifica experimentalment que, en mitjana, els camps locals s'aproximen als valors òptims  $\Delta^0 F^*$  de l'apartat 2.3. Recolzats en aquesta hipòtesi, i per a qualsevol funció simètrica, definim la *temperatura normalitzada* com

$$\hat{T} = \frac{T}{\Delta^0 F^*}. \quad (5.4)$$

Les dues darreres columnes de la figura 5.7 mostren els camps locals, i els valors normalitzats de  $T_c^*$ . Com es pot observar, la dispersió es redueix considerablement. Es important recordar que el valor de  $\Delta^0 F^*$  depèn del disseny i la funció, de manera que la generalització queda implícita. Pel que fa a  $T_c$ , doncs, la definició del valor estàndard es farà en el domini normalitzat.

Els paràmetres estàndard que escollim són

$$\begin{aligned} \hat{T}_c^{std} &= 0.110 \\ N_t^{std} &= 100. \end{aligned} \quad (5.5)$$

L'estratègia resultant la denotarem per

$$RS^{std} \equiv RS(T_c^{std}, 2, N_t^{std}). \quad (5.6)$$

Els casos que més es desvien de la regió òptima són d18 i d21. Però és necessari acceptar un compromís.

La pèrdua de qualitat entre els punts òptim i estàndard s'avalua en el següent experiment.

### Experiment 8

**Objectius:** Comparació entre els resultats *òptims* i *estàndard* de la RS.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ .

**Estratègies:**  $RS^{opt}$ ,  $RS^{std}$ .

**Rèpliques:** 15 per als òptims, 5 per als estàndard.

La taula de la figura 5.8 presenta els resultats. Com es pot comprovar, hi ha poques diferències significatives (el cas més clar és d18). Però en termes relatius les diferències són petites.

El següent experiment verifica l'optimalitat dels paràmetres estàndard, i la seva generalització a l'altra funció finalista  $F_{uthvq}$ .

### Experiment 9

**Objectius:** Validació dels paràmetres estàndard de RS.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ ,  $F_{uthvq}$ .

**Estratègies:**  $RS^{std}$ .

**Rèpliques:** 3.

**Descripció:** Per a cada problema i funció de cost, un experiment factorial amb punt central, a l'entorn dels valors estàndard:  $\hat{T}_c = 0.110 \pm 0.030$ , i  $N_t = 100 \pm 50$ .

Com en la fase d'optimització, però prenent uns increments més amplis, la validació consisteix en comprovar si la resposta al punt estàndard és efectivament millor que la resposta al seu entorn. Els resultats no es mostren en detall. Pel que fa a  $F_{uq}$ , la funció d'entrenament, els 5 primers problemes validen l'optimalitat dels paràmetres estàndard (en el cas de d2, amb  $N_t \gg 1$ , és curiós verificar com els paràmetres estàndard són, localment, els millors). El problema d18 respon millor per a  $T_c$  i  $N_t$  més grans. Mentre que d21 demana  $T_c$  més petita i  $N_t$  més gran. Aquests resultats són consistents amb l'elecció dels paràmetres estàndard, i il·lustren les conseqüències del compromís.

Pel que fa a la segona funció de cost  $F_{uthvq}$ , la distribució de resultats és sorprenentment similar. Amb l'excepció del cas d15, que demana  $T_c$  i  $N_t$  més petits (el punt estàndard és el segon millor), la localització de la millor resposta es repeteix cas a cas. I la qualitat relativa dels punts de l'entorn és molt semblant. La generalització, doncs, sembla molt bona.

Així, pel que fa al conjunt d'entrenament, els paràmetres estàndard escollits són bons.

#### 5.1.4 Paràmetres estàndard per a la CC

El procés d'entrenament de la cerca cooperativa  $CC(S, R_0, R_f, N_t)$  és similar al de la RS. Els paràmetres a optimitzar són el nombre  $S$  de membres de l'esquadra, els radis inicial  $R_0$  i final  $R_f$ , i el nombre total  $N_t$  d'iteracions de la fase cooperativa. A més, en aquest cas, cal tenir en compte les tres variants de la cerca: focalitzada descendent CCfd, focalitzada central CCfc, i dispersa CCd. Donat que es tracta



d'una estratègia nova, no tenim coneixement del procés, i l'exploració s'ha de fer a cegues.

El primer experiment busca els paràmetres òptims per als problemes del conjunt d'entrenament, tractant de manera independent cada una de les variants de l'estratègia.

### Experiment 10

**Objectius:** Optimització dels paràmetres de CC.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ .

**Estratègies:** CCfd, CCfc, CCd.

**Rèpliques:** 1.

**Descripció:** De manera independent per a cada problema i variant de l'estratègia, optimització seqüencial d' $w$  respecte als paràmetres (veure text).

L'optimització es realitza, com abans, a base d'experiments factorials amb punt central, basats en els 4 paràmetres de l'estratègia. I, per tal de desfer ambigüetats, es mostregen uniformement algunes àrees d'interès. Els resultats no recolzen la hipòtesi de treball. En primer lloc les desviacions, molt més grans que en el cas anterior, enmascaren les diferències. I la optimització es fa més difícil. El procés és sovint insensible als radis inicial i final. I, quan es detecta una zona bona, el marge d'optimalitat és bastant ampli. Pel que fa al nombre de membres, la cooperació millora efectivament la freqüència d'èxit de l'esquadra. Però el cost invertit és tan gran que els resultats recomanen esquadres petites ( $S=2$  o  $4$ ). Pel mateix raonament, els òptims de  $N_t$  tendeixen a valors baixos (entre 1 i 20). Com a l'apartat anterior, d2 tendeix a suprimir la fase cooperativa, i no aporta informació d'interès. De manera que l'eliminem del conjunt d'entrenament.

Amb  $S = 2$ , el cas més freqüent, CCfc i CCd són la mateixa estratègia (un únic membre, a més de la referència, no discrimina el tipus de topologia). I la diferència amb CCfd rau en el procés de contracció, que incorpora cerca descendent. D'una variant (CCfd) a l'altra (CCfc-CCd), la localització dels òptims varia sensiblement. Però els resultats òptims obtinguts amb la primera superen, cas a cas, als òptims de les altres dos. De manera que eliminem les dues darreres.

Amb aquests resultats, escollim manualment els els punts *òptims* de treball que es mostren a la taula de la figura 5.9. I l'estratègia corresponent la denotarem per

$$CC^{opt} \equiv CCfd(S^*, R_0^*, R_f^*, N_t^*). \quad (5.7)$$

Pel que fa al radi inicial, la diversitat és gran, no trobem cap relació amb cap paràmetre conegut. Per a d4, d15 i d21,  $R_0 = vb$ , el tamany del disseny. Però en els altres tres casos aquesta elecció empitjora els resultats. Pel que fa al

nombre d'iteracions, malgrat que no s'observa clarament en els òptims escollits, l'experimentació realitzada recomana un  $N_t$  més gran per als problemes més grans. A falta de res millor, doncs, obtem per pendre els valors següents com a paràmetres estàndard:

$$\begin{aligned} S^{std} &= 2 \\ R_0^{std} &= vb \\ R_f^{std} &= 4 \\ N_t^{std} &= 15, \end{aligned} \tag{5.8}$$

que configuren la cerca cooperativa estàndard

$$CC^{std} \equiv CCfd(S^{std}, R_0^{std}, R_f^{std}, N_t^{std}). \tag{5.9}$$

El següent experiment compara els resultats al punt òptim i els resultats al punt estàndard.

#### Experiment 11

**Objectius:** Comparació entre els resultats *òptims* i *estàndard* de la CC.

**Problemes:** Conjunt d'entrenament, menys d2.

**Funcions:**  $F_{uq}$ .

**Estratègies:**  $CC^{opt}$ ,  $CC^{std}$ .

**Rèpliques:** 15 per als òptims, i 5 per als estàndard.

La taula de la figura 5.10 mostra els resultats. Les diferències, en aquest cas, són sempre significatives, i les diferències són grans.

Per a la CC, doncs, l'estandardització, que no és localment òptima, perjudica fortament els resultats.

#### 5.1.5 Altres proves

L'algorisme bàsic de cerca del capítol 3, que hem analitzat fins ara, inclou algunes decisions implícites que no s'havien formulat com variants. La primera, apuntada a peu de pàgina al final de l'apartat 3.1, fa referència a l'*ordre* d'actualització de les unitats de la xarxa. L'ordre proposat és *seqüencial* (una unitat darrera l'altre, ordenades per files), amb la intenció de reduir tota l'aleatorietat a l'elecció de l'estat inicial. L'altra alternativa és escollir l'ordre *a l'atzar*. Els experiments preliminars, realitzats amb CD i RS, mostren que les diferències són poc significatives (amb un lleuger avantatge per a l'ordre a l'atzar). Però, en la implementació paral·lela, la segona alternativa és computacionalment més costosa, i es va prescindir de l'actualització a l'atzar.

La nota a peu de pàgina de l'apartat 3.2 apunta un altre fet interessant. La regla de decisió, en actualitzar la unitat  $(i, j)$ , depèn del signe de l'increment de

transició  $\Delta_{transc}^{ij}E$ . Però, quan aquest increment és zero (un cas molt freqüent amb les nostres funcions de cost), les alternatives són tres: afavorir els *zeros* (commutar tans sols si  $x_{ij} = 1$ ), afavorir els *uns* (commutar quan  $x_{ij} = 0$ ), i no afavorir a *ningú* (no commutar). Obviament, l'alternativa de commutar sempre no és possible, ja que podria provocar cicles. Els experiments preliminars amb CD van demostrar diferències importants. La opció de no afavorir a ningú era notablement més dolenta, i la d'afavorir els zeros era lleugerament millor que els uns. De manera que els experiments d'aquest capítol s'han fet amb aquesta alternativa. Per extensió, s'ha aplicat el mateix criteri a les altres estratègies.

Finalment, la darrera puntualització fa referència al conjunt de problemes. Al llistat de [Mathon & Rosa, 90] apareixen només els dissenys amb  $u \leq vb/2$ . Els paràmetres admissibles amb  $u > vb$  queden implícits, doncs corresponen a complementaris d'algun disseny de la llista. La darrera prova que es va fer consisteix en generar aquests problemes amb nombre d'uns elevat (generació pel disseny *complementari*). El fracàs va ser sorprenent. Dels 5 dissenys del conjunt d'entrenament amb  $u < vb$  (d4, d8, d10, d15, d21), únicament d4 no va donar fracàs total amb CD. El resultat sembla apuntar que l'algorisme de cerca es comporta millor quan la densitat d'uns és baixa.

## 5.2 Segona etapa: Comparacions

Les dades de l'apartat anterior es recopilen i completen en els dos experiments següents:

### Experiment 12

**Objectius:** Recopilació dels resultats òptims.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ .

**Estratègies:** CD,  $RS^{opt}$ ,  $CC^{opt}$ .

**Rèpliques:** 15.

**Comentaris:** Dades extretes dels experiments 6 (CD), 8 (RS) i 11 (CC).

### Experiment 13

**Objectius:** Recopilació dels resultats estàndard, per a les *dues* funcions finalistes.

**Problemes:** Conjunt d'entrenament.

**Funcions:**  $F_{uq}$ ,  $F_{uthvq}$ .

**Estratègies:** CD,  $RS^{std}$ ,  $CC^{std}$ .

**Rèpliques:** 5.

**Comentaris:** Recopilació de diversos experiments i extensió a la segona funció (en el cas d'RS i CC).

L'experimentació s'exten, en aquest apartat, als 18 problemes que falten per a completar els 25 del conjunt de prova.

#### Experiment 14

**Objectius:** Avaluació dels 18 problemes restants del conjunt de prova.

**Problemes:** Conjunt de prova.

**Funcions:**  $F_{uq}$ ,  $F_{uthvq}$ .

**Estratègies:** CD, RS<sup>std</sup>, CC<sup>std</sup>.

**Rèpliques:** 5.

Els resultats s'analitzen a continuació per a cada un dels factors.

#### 5.2.1 Selecció de la millor funció per a cada estratègia

Donat que volem determinar el millor *algorithme*, el que de debò necessitem és escollir la funció més adequada a cada estratègia.

A la taula de la figura 5.11, es comparen, estratègia per estratègia, els resultats de les dues funcions finalistes en els experiments anteriors. El cas més difícil de discernir correspon a CD. En total,  $F_{uq}$  resol un cas de menys (el cas d8, ja que d21 i d6 es compensen). Però, dels 12 problemes que resolen totes dues, en 8 ocasions els resultats d' $F_{uq}$  són significativament més bons. Mentre que els d' $F_{uthvq}$  només ho són en dues. La composició de la darrera, apresada al conjunt d'entrenament, resulta inferior en el conjunt de prova (exceptuant d6). Tenint en compte, a més, que els casos que guanya  $F_{uq}$  són els dissenys més grans, l'escollim com a vencedora. En el cas de la RS,  $F_{uq}$  és qui resol un cas de més (d19, tot i que amb dificultat). I dels 23 casos que resolen totes dues, en 15  $F_{uq}$  és significativament millor, mentre que  $F_{uthvq}$  només ho és en 5. La conclusió, doncs, és clara. I finalment, pel que fa a CC,  $F_{uq}$  resol novament dos casos més (d16 i d17, també amb dificultat). I dels 20 problemes resolts per totes dues, les diferències significatives són de 14 a 2. Algunes d'elles, a favor d' $F_{uq}$ , molt clares (d15, d11 i d25). L'entrenament d' $F_{uthvq}$ , fet amb CD, no generalitza gaire bé a les altres estratègies.

La conclusió general, doncs, és clara. I prenem  $F_{uq}$  com a funció guanyadora per a les tres estratègies. Els avantatges d'aquest resultat són que es tracta de la funció d'entrenament (amb la qual s'han fixat els paràmetres), i que la seva composició, al seu torn, no ha estat apresada (els seus coeficients són fixos).

#### 5.2.2 Comparació entre estratègies

Seguint el fil de l'apartat anterior, la comparació entre estratègies es fa amb la millor funció de cada cas. Es a dir, es fa sempre amb  $F_{uq}$ . En termes general, però, les

conclusions que obtindrem apliquen igualment als resultats obtinguts amb  $F_{uthvq}$ , que no s'analitzen en detall. Formalment, les dades corresponents es recopilen en el següent experiment.

#### Experiment 15

**Objectius:** Comparació entre les millors estratègies.

**Problemes:** Conjunt d'entrenament complet.

**Funcions:**  $F_{uq}$ .

**Estratègies:** CD,  $RS^{std}$ ,  $CC^{std}$ .

**Rèpliques:** 5.

**Comentaris:** Dades procedents dels experiments 13 i 14.

Els resultats de la comparació es mostren a la taula de la figura 5.12 que en cada cas, quan les diferències són significatives, dictamina la primera i la segona posició. La conclusió general és clara. Cas a cas, la recuita simulada supera a les altres dues, amb diferències importants i significatives. L'única excepció és d2, comentat anteriorment, que és tan fàcil per a CD que RS aconseguix, tot just, igualar-la. L'avantatge es manifesta tan en el nombre de problemes resolts (tots menys d22), com en les diferències entre resultats.

La comparació entre les estratègies restants és interessant perquè medeix els efectes de la cooperació en CC respecte de la cerca descendent. El fenomen que s'observa és el següent. Dels 13 casos que resolen totes dues, CD és significativament superior en 9 i CC només en 2. Però CC és capaç de trobar solució a 9 dissenys en els que CD fracassa. Això ve a demostrar que CC augmenta considerablement la probabilitat d'èxit, però està fortament castigada pel seu cost addicional.

Pel que fa al conjunt d'entrenament, la comparació feta amb paràmetres òptims es mostra a la figura 5.13). CC aconseguix una victòria a d18, i el marcador entre CD i CC queda 0 a 3, sobre 4 casos comuns, i 2 casos que únicament CC resol. Les conseqüències de l'estandardització són ara manifestes.

### 5.2.3 Validació dels paràmetres estàndard de la RS

Com s'ha vist a l'apartat 5.1.3, la generalització a  $F_{uthvq}$  dels paràmetres estàndard de la recuita simulada és bona. El següent experiment verifica a continuació la generalització d'aquests paràmetres als problemes del conjunt de prova.

#### Experiment 16

**Objectius:** Validació dels paràmetres estàndard de RS.

**Problemes:** Conjunt de prova.

**Funcions:**  $F_{uq}$ .

**Estratègies:**  $RS^{std}$ .

**Rèpliques:** 1.

**Descripció:** Per a cada problema i funció de cost, un experiment factorial amb punt central, a l'entorn dels valors estàndard:  $\hat{T}_c = 0.110 \pm 0.030$ , i  $N_t = 100 \pm 50$ .

Els resultats, que no es mostren en detall, es descriuen a continuació: Si ordenem, en cada cas, les 5 respostes obtingudes (punt central i quatre extrems), dels 18 problemes avaluats el punt central queda 11 cops en primera posició, sis d'ells amb avantatge clara (d7, d9, d11, d14, d20 i d25). Queda 5 cops en segona posició (tres d'ells per poca diferència). I 1 cop en tercera posició (d13), poc allunyat de la segona. El cas restant, d22, queda irresolt. Fora del punt central, les segones (o primeres) posicions apareixen amb més freqüència a l'extrem amb  $\hat{T}_c$  gran i  $N_t$  petit. En termes generals, però, l'efecte del primer paràmetre torna a ser més important.

En definitiva, la generalització respecte dels problemes també és bona.

#### 5.2.4 Anàlisi dels resultats en funció del tamany del problema

Fetes les comparacions, les dades de l'experiment 15 s'analitzen a continuació en termes de la variable *resultat*  $z$ , i s'observa la seva evolució amb el tamany  $vb$  del problema. Per mitjà de l'experiment paral·lel, els paràmetres estimats són  $\hat{E}(z)$  i  $\widehat{Dev}(z)$ . Però, com hem fet a l'apartat 3.1, analitzarem prèviament els paràmetres que intervenen en aquestes estimacions.

Abans que res, però, parlem del temps d'execució. La figura 5.14 mostra la mitjana del temps d'iteració en funció del tamany al quadrat del problema. Les dades corresponen a 1 descens d'una execució a l'entorn descrit al capítol 4 (temps de procés), i constaten l'anàlisi de complexitat de l'apartat 4.1.1. La variabilitat que s'observa en els gràfics depèn dels resultats de la cerca i de la variabilitat intrínseca de les eines de mesura. Però les corbes obtingudes són essencialment lineals (complexitat  $\mathcal{O}(v^2b^2)$ ). El temps d'iteració, com predèiem, és més gran per a les estratègies més complexes ( $CC$  i  $RS$ , en aquest ordre). Però les diferències són poc importants. El temps total d'execució d'una rèplica de l'experiment paral·lel, per als 25 problemes i amb paràmetres estàndard, és d'1h 35' per a  $CD$ , 10h 11' per a  $RS$ , i 5h 2' per a  $CC$ . L'experiment no acaba fins que l'últim processador convergeix, i això succeix al cap d'unes 10 iteracions aproximadament per a  $CD$ , 100 per a  $RS$ , i 40 per a  $CC$ . El temps total d'execució de l'experiment 15 és, doncs, de 3 dies i 9 hores. Per a  $RS$  (la millor estratègia en termes de resultats), estimant el pendent del gràfic en  $3.5 \times 10^{-4}$ , el tamany *màxim* d'un problema que es pugui executar en 48 hores (1 descens) és, sent optimistes,  $vb = 2500$ .

El següent terme que s'analitza és el cost de relaxació (en iteracions) d'un descens. Distingint el cost  $b$  dels acabats en èxit del cost  $d$  dels acabats en fracàs. Les dades de l'experiment 15 ens proporcionen, en cada cas, un total de  $5 \times 2048$  descensos,

i els resultats, en mitjana, es mostren a la figura 5.15 en funció del tamany  $vb$  del problema. Donat que el nombre d'èxits és variable, el nombre de mostres varia i les mitjanes que es presenten no tenen la mateixa precisió (en els casos de fracàs total,  $b$  simplement no es mostra). Però, en la major part dels casos el nombre d'observacions és tan gran, que les diferències entre  $\bar{b}$  i  $\bar{d}$  que la figura permet distingir, són estadísticament significatives. Això corrobora els models de l'apartat 4.1.2.

La impressió general és que el cost  $b$  dels descens bons, amb una evolució una mica erràtica, tendeix a créixer lleugerament amb el tamany del problema, i l'evolució del cost  $d$  depèn fortament de l'estratègia. En el cas de CD les dues corbes es segueixen (el pic de  $b$  a  $vb = 256$ , corresponent a  $d_{21}$ , no és significatiu, ja que únicament s'han trobat dues solucions), i la tendència al creixement és molt lleugera. Les diferències, tot i significatives, són petites, amb el cost dels descens bons lleugerament per sobre. Les desviacions, que no es mostren, són també del mateix ordre (entre 0.7 i 1.5, en ambdós casos).

Pel que fa a RS, el valor de  $\bar{b}$  és sempre sensiblement inferior a  $N_t = 100$  (amb desviacions de l'ordre de 20 iteracions), i el valor de  $\bar{d}$  no supera mai 101.5 (amb desviacions inferiors a 0.8). El comportament és, doncs, el que hem descrit a l'apartat 5.1. Els òptims apareixen en la fase estocàstica de la cerca, i els mínims locals acaben tot just superat  $N_t$  (la relaxació final amb CD és gairabé sempre innecessària). Amb aquesta estratègia, el creixement de  $\bar{b}$  amb  $vb$  és molt marcat. Donat, a més, que les desviacions són força uniformes, aquest fet suggereix que un creixement lineal de  $N_t$  amb  $vb$  podria millorar els resultats. La complexitat global d'un descens seria, aleshores,  $\mathcal{O}(v^3b^3)$ .

Finalment, en el cas de la CC, es confirma novament l'anàlisi de l'apartat 5.1. Òptims i mínims locals apareixen gairabé sempre després de la fase cooperativa, i la relaxació amb CD és, en els dos casos, necessària. La corba dels mínims locals segueix molt de prop la de la CD, desplaçada  $S \times N_t = 30$  iteracions, amb desviacions similars. Però la dels òptims és fa més erràtica, amb desviacions més grans (entre 1.5 i 4). El cost dels descens bons, en aquest cas, tendeix a ser inferior al dels locals.

Les taules de les figures 5.16, 5.17, i 5.18 mostren les estimacions  $\hat{E}(y)$ ,  $\widehat{Dev}(y)$ ,  $\hat{E}(z)$  i  $\widehat{Dev}(z)$ , per a cada una de les estratègies (mitjana i desviació, en cada cas, de les 5 rèpliques). I els gràfics que venen a continuació (figures 5.19 i 5.20) mostren l'evolució de les esperances estimades amb el tamany  $vb$  del problema. Donat que les mitjanes de les taules no ponderen els casos de fracàs complet (el cost de les rèpliques fracassades no es té en compte), la seva comparació gràfica sortiria enmascarada, de manera que hem optat per representar la reinterpretació de les dades com un únic experiment paral·lel amb  $N = 5 \times 2048 = 10240$  processadors. Els casos no resolts, però, s'omiteixen.

El nombre d'intents  $y$  fins la primera solució ens parla de l'eficàcia de la cerca, i el seu valor estimat varia considerablement amb el problema i l'estratègia. Les desviacions estimades, similars a les mitjanes, no aporten informació d'interès (com correspon a una distribució geomètrica, la desviació és la mitjana ponderada per  $\sqrt{q}$ , amb  $q$  la probabilitat de fracàs). La precisió de les estimacions millora quan el valor

és petit. La figura 5.19 a) mostra l'evolució d' $\hat{E}(y)$ . Però la representació gràfica resulta molt més clara en termes de l'estimador invers: la probabilitat estimada  $\hat{p}$  (freqüència d'èxit), que es mostra a la figura 5.19 b). Qualitativament, la forma de les curves és essencialment paral·lela, de manera que la dificultat relativa dels problemes es manté d'una estratègia a una altra. Si bé l'evolució local és molt erràtica, la dificultat dels problemes tendeix a augmentar amb el tamany. La superioritat d'RS es veu clara. Però el més important de remarcar, com es pot verificar a les taules, és que CC supera cas a cas a CD. En termes d'eficàcia, l'ús d'esquadres més grans podria millorar els seus resultats.

El cost  $z$  fins la primera solució, el resultat principal de l'anàlisi, medeix l'eficiència de la cerca, i n'estimem l'esperança i la variança. Per inspecció de les taules, i en termes de desviació, la relació entre aquest paràmetres és similar al cas d' $y$ . Ambdós valors són similars, amb diferències més petites conforma augmenta el valor absolut. Però la precisió de les mesures, com hem vist a l'apartat 5.1, empitjora. Com abans, l'evolució d' $\hat{E}(z)$  amb el tamany del problema (figura 5.20 a)) és il·legible. De manera que la figura 5.20 b) (ampliada a la figura 5.20 c)) en mostra una inversa. La variable que es presenta,  $\frac{10.000}{w}$ , no té una interpretació física estricta, però és una estimació "informal" del nombre d'òptims que obtindriem en 10.000 iteracions. En consonància amb el que acabem de dir, l'evolució de les corbes és similar a les de la freqüència d'èxit. A mesura que creix  $vb$  les dificultats augmenten, però cada problema mostra el seu comportament particular. Si comparem les estratègies, en termes d'eficiència el baix cost de CD escurça les distàncies amb RS, i arriba, com hem vist, a superar a CC en la major part dels casos que totes dues resolen.

### 5.3 Tercera etapa: Aplicació del millor algorisme a problemes de tamany creixent

El millor algorisme obtingut (RS<sup>std</sup> amb  $F_{uq}$ ) s'aplica en aquest apartat a problemes de tamany creixent, seguint la taula 4.1. El primer experiment avalua els 32 casos següents (d26 a d57), amb  $vb \leq 500$ , però es deixen de banda d27 i d55, dels quals sabem que no existeixen.

#### Experiment 17

**Objectius:** Anàlisi de la resposta per a dissenys de tamany creixent.

**Problemes:** Disseny solubles amb  $vb \leq 500$  (30 problemes).

**Funcions:**  $F_{uq}$ .

**Estratègies:** RS<sup>std</sup>.

**Rèpliques:** 1.



Els resultats es mostren a la taula de la figura 5.21. Dels 30 casos intentats, 14 queden irresolts, i els 16 casos restants mostren valors molt elevats d' $\hat{E}(\mathbf{z})$ . No s'observa una proporció directa amb  $vb$ , però es fa evident que, a mesura que creix el tamany, les dificultats per resoldre problemes augmenten.

Com hem vist, amb  $vb$  gran, el temps d'execució de l'experiment paral·lel és elevat. De manera que, a continuació, optem per abortar el procés en el moment que algun processador troba la primera solució. Així, el nombre total  $N_t$  d'iteracions només s'esgotarà en els casos de fracàs total. Els 72 casos que s'avaluen al següent experiment (d58 a d129), completen la llista de dissenys amb  $vb \leq 1000$ . I novament eliminem d71 i d105 perquè no tenen solució.

### Experiment 18

**Objectius:** Generació de dissenys de tamany creixent.

**Problemes:** Dissenys solubles amb  $vb \leq 1000$  (70 problemes).

**Funcions:**  $F_{uq}$ .

**Estratègies:**  $RS^{std}$ .

**Rèpliques:** 1.

**Descripció:** Experiment paral·lel truncat a la primera solució (no cal acabar la resta de descensos). Si no en trobem cap, tenim un fracàs total (cap disseny en  $N = 2048$  intents).

Els resultats es mostren a la taula de la figura 5.22, que parla per si sola. El disseny més gran resolt té  $vb = 961$ . Però dels 70 problemes avaluats, únicament se n'han pogut resoldre 8.

En conclusió, tal com era d'esperar tenint en compte la complexitat del cas general (NP-complet), la qualitat de l'algorisme de cerca decau per a problemes grans, fins que és incapaç de resoldre'ls. El que s'ha vist en aquest apartat és que aquest tamany *crític* es pot establir pels vols de  $vb \simeq 500$ .

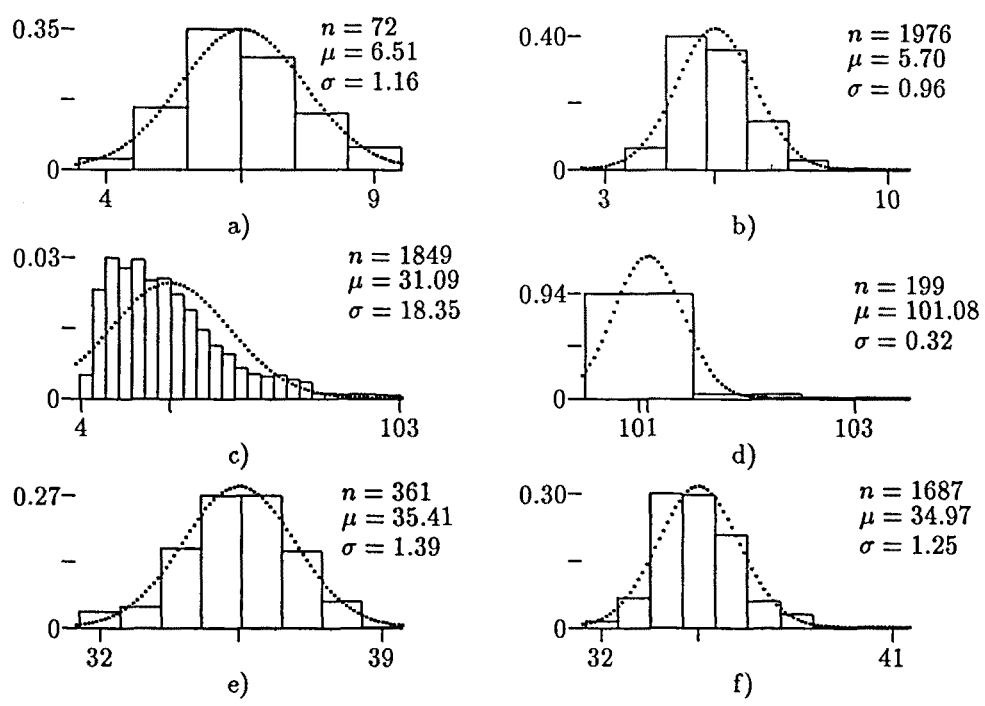


Figura 5.1: (Experiment 1) Per a cada estratègia, histograma del cost  $b$  dels descensos acabats en èxit (esquerra), i del cost  $d$  dels descensos acabats en fracàs (dreta). De dalt a baix, CD, RS i CC. ( $n$ , nombre de mostres).

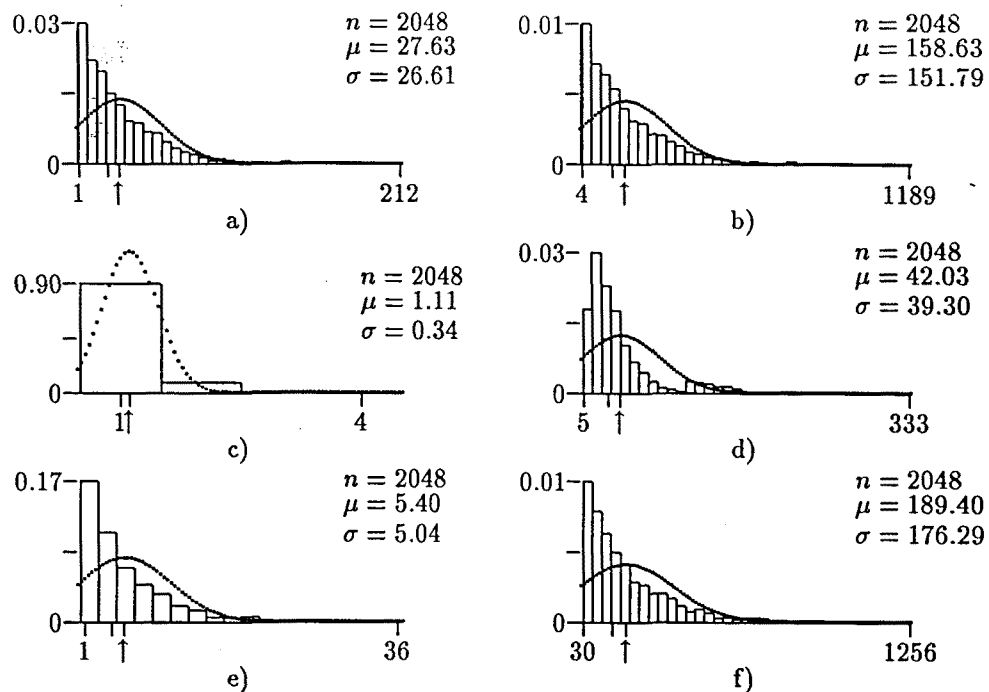


Figura 5.2: (Experiment 2) Per a cada estratègia, els gràfics de l'esquerra mostren l'histograma del nombre de descensos  $y$  fins la primera solució, i els gràfics de la dreta l'histograma del cost  $z$  corresponent. De dalt a baix, CD, RS i CC. A l'eix d'abscises, entre els valors mínim i màxim, el signe  $\uparrow$  assenyalava la mitjana i l'altra indicació correspon a la mediana.

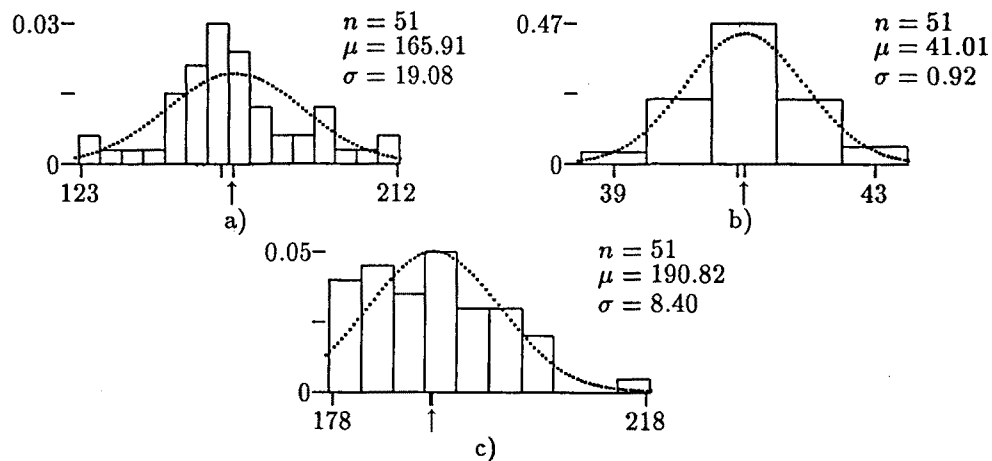


Figura 5.3: (Experiment 3) Per a cada estratègia, histograma de la resposta  $w$ . a) CD, b) RS i c) CC.

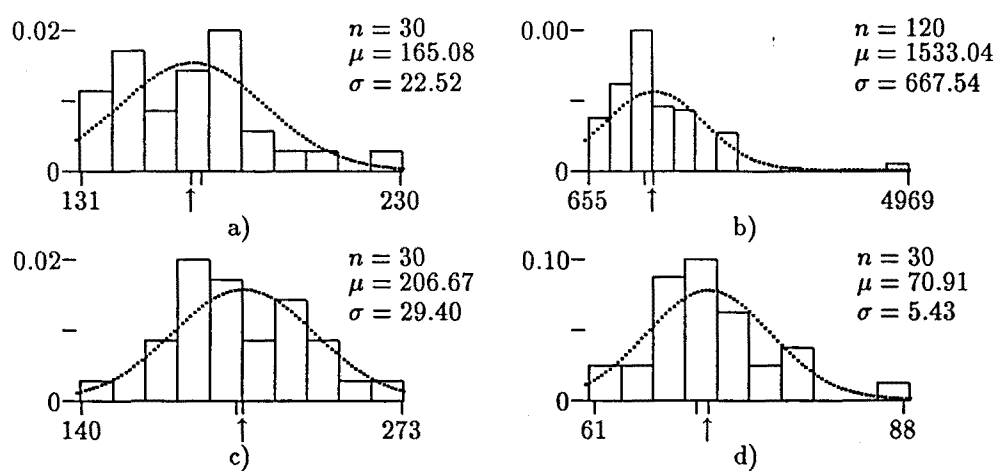


Figura 5.4: (Experiment 4) Histograma de la resposta  $w$  amb CD, per a a) d5, b) d10, c) d7 i d) d11.

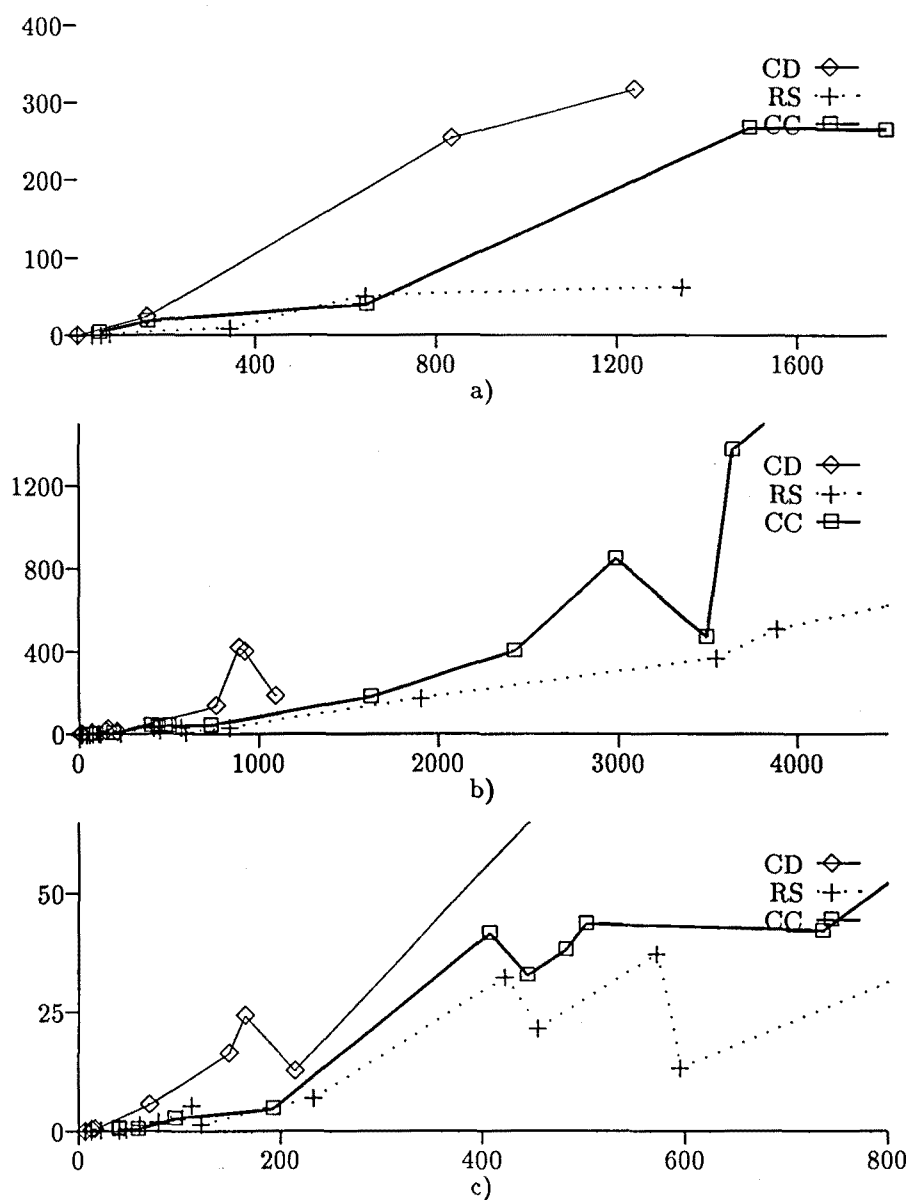


Figura 5.5: (Experiments 12 i 15) Varibilitat de  $w$  (desviacions respecte mitjanes) per a cada una de les estratègies. a) Paràmetres òptims i conjunt d'entrenament (15 rèpliques). b) Paràmetres estàndard i conjunts d'entrenament i prova (5 rèpliques). c) Ampliació de l'anterior.

	$F_{uq}$	$F_{uhvq}$	$F_{uthq}$	$F_{uthvq}$
d2	★ 7.3 (0.1)	11.2 (0.4)	30.1 (1.4)	28.1 (1.6)
d4	837.4 (254.9)	576.5 (160.6)	212.6 (28.2)	○ 138.0 (13.4)
d8	-	-	487.9 (95.0)	458.2 (224.0)
d10	★ 1241.9 (318.4)	[14] 6958.5 (2602.0)	1070.1 (456.4)	1033.9 (282.7)
d15	★ 162.4 (24.4)	232.7 (37.5)	385.4 (63.5)	451.4 (84.3)
d18	-	-	-	-
d21	★ [7] 11095 (3156)	[6] 12926 (81.9)	[1] 15561	[1] 12741

Figura 5.6: (Experiment 6) Resposta  $w$  de les 4 funcions prefinalistes, per a cada problema del conjunt d’entrenament (mitjana i desviació, entre parèntesi, de les 15 rèpliques de l’experiment). Si alguna de les rèpliques és un fracàs complet s’exclou de la mitjana, i el valor que mostren les claus indica el nombre de cassos ponderats. Si totes 15 fracassen (problema irresolt), s’assenyala amb un guió. El símbol “★ ” indica la millor resposta entre  $F_{uq}$  i  $F_{uhvq}$  (si la diferència és significativa). I “○ ” s’utilitza per comparar les altres dues funcions.

	$T_c^*$	$N_t^*$	$\Delta^0 F^*$	$\bar{T}_c^*$
d2	-	1	4	-
d4	0.665	120	6	0.111
d8	1.012	80	10	0.101
d10	0.677	45	6	0.113
d15	0.419	130	4	0.105
d18	1.250	180	10	0.125
d21	1.292	170	14	0.092

Figura 5.7: (Experiment 7) Paràmetres òptims de la RS per al conjunt d’entrenament, amb  $F_{uq}$ . Les dos darreres columnes mostren el camp local dels òptims i la normalització de  $T_c^*$  respecte d’aquests valors. En el cas de d2 els resultats obtinguts no depenen de  $T_c$  (veure text).

	Optims	Estàndard
d2	★ 7.24 (0.17)	7.90 (0.15)
d4	78.98 (2.15)	80.04 (1.74)
d8	59.39 (0.94)	61.90 (1.39)
d10	644.63 (52.02)	572.44 (37.51)
d15	★ 39.48 (1.05)	41.15 (0.25)
d18	★ 1345.87 (63.03)	1906.80 (174.95)
d21	★ 345.68 (9.55)	421.95 (32.56)

Figura 5.8: (Experiment 8) Comparació entre la resposta de RS amb paràmetres òptims (15 rèpliques) i amb paràmetres estàndard (5 rèpliques), amb  $F_{uq}$ .

	$vb$	$S^*$	$R_0^*$	$R_f^*$	$N_t^*$
d4	108	2	108	5	1
d8	147	4	35	3	7
d10	162	4	15	7	1
d15	216	2	216	7	1
d18	240	2	90	1	20
d21	256	2	256	3	10

Figura 5.9: (Experiment 10) Paràmetres òptims de la CC per al conjunt d'entrenament, amb  $F_{uq}$ . La segona columna mostra el tamany  $vb$  del disseny.

	Optims		Estàndard	
d4	* 163.80	(18.39)	407.79	(41.71)
d8	* 1793.63	(266.01)	3641.94	(1374.37)
d10	* 1496.15	(268.28)	3497.37	(472.69)
d15	* 57.30	(4.00)	192.99	(4.75)
d18	* 646.61	(40.53)	2988.41	(848.06)
d21	* [10] 30373.63	(13519.10)	[4] 43322.58	(21372.46)

Figura 5.10: (Experiment 11) Comparació entre els resultats de la CC amb paràmetres òptims (15 rèpliques) i amb paràmetres estàndard (5 rèpliques), amb  $F_{uq}$ .

	CD		RS		CC	
	$F_{uq}$	$F_{uthvq}$	$F_{uq}$	$F_{uthvq}$	$F_{uq}$	$F_{uthvq}$
d2	★ 7.3	28.2	★ 7.9	23.6	★ 41.3	116.5
d4	926.6	★ 133.8	★ 80.0	94.2	407.8	353.9
d8	-	★ 590.0	★ 61.9	86.5	3641.9	★ 1395.5
d10	1095.5	917.3	★ 572.4	2207.2	3497.4	3504.0
d15	★ 164.8	504.5	★ 41.1	195.8	★ 193.0	1230.4
d18	-	-	1906.8	636.3	★ 2988.4	5148.0
d21	★ [1] 6443.5	-	★ 421.9	10182.6	★ [4] 43322.6	[1] 74145.0
d1	13.2	★ 9.8	★ 7.3	10.3	★ 40.5	47.9
d3	★ 16.3	49.2	★ 13.0	47.5	★ 59.8	182.8
d5	149.4	117.0	★ 112.3	424.9	482.2	443.5
d6	-	★ [3] 4625.8	122.1	★ 74.2	★ 737.2	1301.9
d7	★ 215.5	1517.1	★ 48.2	326.4	★ 502.5	3396.2
d9	★ 765.1	5277.2	★ 233.9	1494.6	★ 2424.9	5626.0
d11	★ 71.2	722.0	★ 22.9	65.9	★ 95.8	1329.6
d12	★ [3] 9328.0	[1] 10053.0	★ 29441.2	124164.7	[1] 70806.0	★ [3] 72580.0
d13	-	-	453.7	★ 227.1	★ 1630.3	2674.0
d14	-	-	596.0	★ 218.7	★ 9078.2	17342.1
d16	-	-	9381.1	13286.8	★ [1] 34308.0	-
d17	-	-	★ 3894.1	115327.5	★ [1] 73872.0	-
d19	-	-	★ [1] 206995.0	-	-	-
d20	-	-	840.3	★ 567.0	25615.7	36985.5
d22	-	-	-	-	-	-
d23	-	-	3550.3	★ 1156.3	★ 42187.6	[3] 68622.7
d24	-	-	35200.7	45208.2	-	-
d25	★ 892.3	2340.3	★ 79.9	416.1	★ 443.5	4869.5

Figura 5.11: (Experiments 13 i 14) Per a cada una de les estratègies estàndard, comparació entre les dues funcions finalistes (5 rèpliques). Conjunts d'entrenament (quadre superior) i prova (quadre inferior). Per raons d'espai, no es mostren les desviacions, però s'han tingut en compte a l'hora de senyalar les diferències significatives.



	CD		RS		CC	
d2	★ 7.3	(0.1)	○ 7.9	(0.2)	41.3	(0.5)
d4	926.6	(402.6)	★ 80.0	(1.7)	○ 407.8	(41.7)
d8	-		★ 61.9	(1.4)	○ 3641.9	(1374.4)
d10	○ 1095.5	(184.9)	★ 572.4	(37.5)	3497.4	(472.7)
d15	164.8	(24.4)	★ 41.1	(0.2)	193.0	(4.8)
d18	-		★ 1906.8	(175.0)	○ 2988.4	(848.1)
d21	[1] 6443.5		★ 421.9	(32.6)	○ [4] 43322.6	(21372.5)
d1	○ 13.2	(0.4)	★ 7.3	(0.1)	40.5	(0.7)
d3	○ 16.3	(0.6)	★ 13.0	(0.1)	59.8	(0.4)
d5	○ 149.4	(16.5)	★ 112.3	(5.6)	482.2	(38.2)
d6	-		★ 122.1	(1.6)	○ 737.2	(42.2)
d7	○ 215.5	(12.7)	★ 48.2	(0.6)	502.5	(43.6)
d9	○ 765.1	(137.8)	★ 233.9	(7.1)	2424.9	(403.5)
d11	○ 71.2	(5.6)	★ 22.9	(0.4)	95.8	(2.6)
d12	○ [3] 9328.0	(93.8)	★ 29441.2	(5128.6)	[1] 70806.0	
d13	-		★ 453.7	(21.7)	○ 1630.3	(183.0)
d14	-		★ 596.0	(13.5)	○ 9078.2	(5200.2)
d16	-		★ 9381.1	(1507.9)	○ [1] 34308.0	
d17	-		★ 3894.1	(512.9)	○ [1] 73872.0	
d19	-		★ [1] 206995.0		-	
d20	-		★ 840.3	(34.9)	○ 25615.7	(8779.6)
d22	-		-		-	
d23	-		★ 3550.3	(372.2)	○ 42187.6	(24720.7)
d24	-		★ 35200.7	(21245.0)	-	
d25	892.3	(418.6)	★ 79.9	(2.1)	443.5	(32.8)

Figura 5.12: (Experiment 15) Comparació entre estratègies amb la millor funció,  $F_{uq}$ , i paràmetres estàndard. (5 rèpliques). Quan les diferències són significatives, el símbol “★” assenyala la millor resposta, i el símbol “○” la segona.

	CD		RS		CC	
d2	7.3	(0.1)	7.2	(0.2)		
d4	837.4	(254.9)	★ 79.0	(2.2)	○ 163.8	(18.4)
d8	-		★ 59.4	(0.9)	○ 1793.6	(266.0)
d10	1241.9	(318.4)	★ 644.6	(52.0)	1496.2	(268.3)
d15	162.4	(24.4)	★ 39.5	(1.0)	○ 57.3	(4.0)
d18	-		○ 1345.9	(63.0)	★ 646.6	(40.5)
d21	[7] 11094.8	(3155.9)	★ 345.7	(9.6)	○ [10] 30373.6	(13519.1)

Figura 5.13: (Experiment 12) Comparació entre estratègies, amb  $F_{uq}$  i paràmetres òptims, per al conjunt d'entrenament (15 rèpliques). EL cas d2 amb CC no ha estat optimitzat.

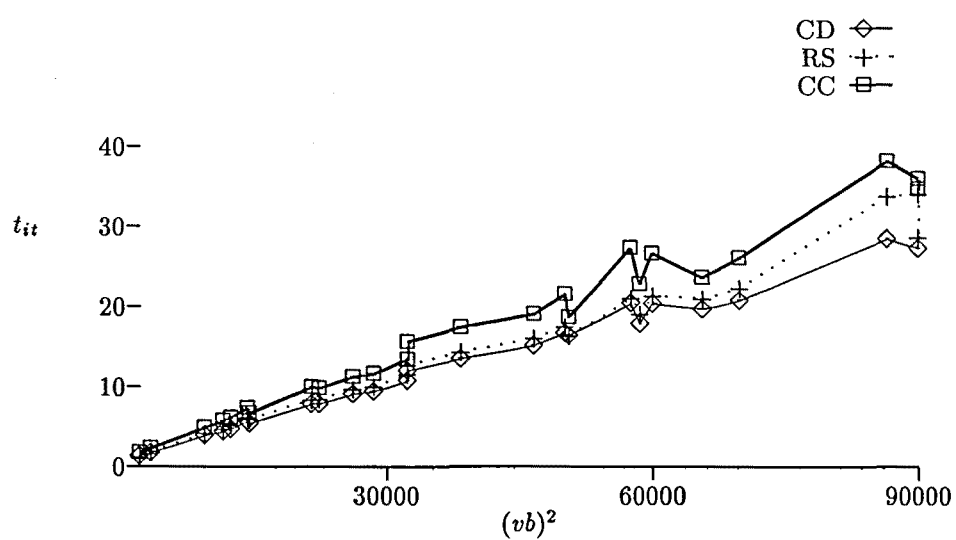


Figura 5.14: (Experiment 15) Temps d'iteració  $t_{it}$  (seg) en funció de  $(vb)^2$ , per a les tres estratègies.

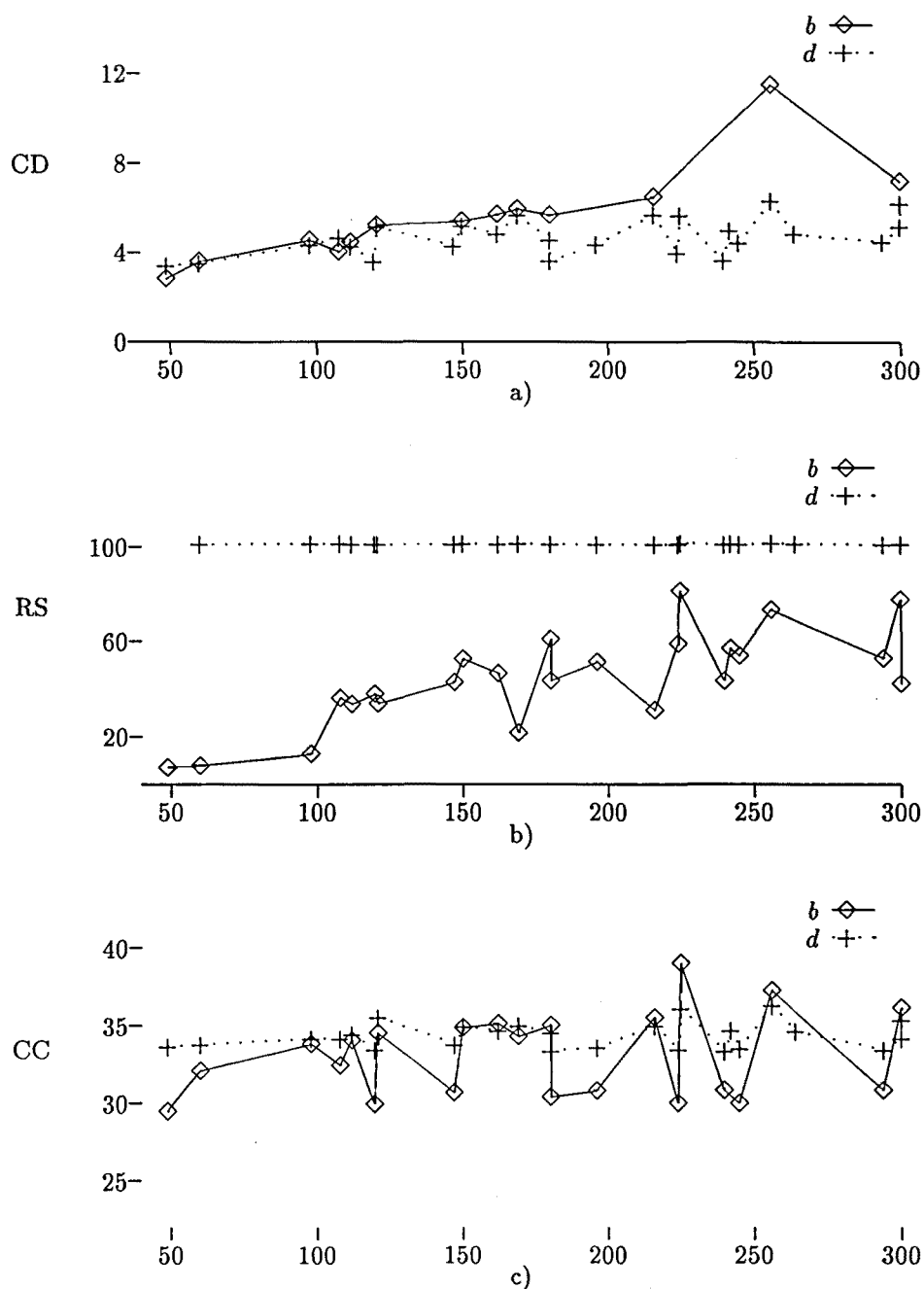


Figura 5.15: (Experiment 15) Evolució del cost  $b$  dels descensos bons i  $d$  dels descensos dolents amb el tamany  $v_b$  del problema. Per a a) CD, b) RS i c) CC. (Sobre un total  $n_b + n_d = 5 \times 2048 = 10240$  descensos, en cada cas).

CD								
	$\hat{E}(y)$		$\widehat{Dev}(y)$		$\hat{E}(z)$		$\widehat{Dev}(z)$	
d1	4.1	(0.1)	3.5	(0.1)	13.2	(0.4)	12.1	(0.4)
d2	2.1	(0.0)	1.5	(0.0)	7.3	(0.1)	5.3	(0.1)
d3	3.7	(0.1)	3.2	(0.1)	16.3	(0.6)	13.8	(0.5)
d4	199.4	(85.7)	198.9	(85.7)	926.6	(402.6)	925.0	(402.7)
d5	35.0	(3.9)	34.4	(3.9)	149.4	(16.5)	147.1	(16.5)
d6	-		-		-		-	
d7	41.9	(2.3)	41.4	(2.3)	215.5	(12.7)	212.9	(12.7)
d8	-		-		-		-	
d9	148.1	(26.7)	147.6	(26.7)	765.1	(137.8)	762.4	(137.6)
d10	226.9	(38.0)	226.4	(38.0)	1095.5	(184.9)	1092.4	(185.4)
d11	12.5	(1.0)	12.0	(1.0)	71.2	(5.6)	68.1	(5.6)
d12	[3] 2048.0	(0.0)	[3] 2047.5	(0.0)	[3] 9328.0	(93.8)	[3] 9324.7	(93.4)
d13	-		-		-		-	
d14	-		-		-		-	
d15	28.8	(4.3)	28.3	(4.3)	164.8	(24.4)	161.2	(24.4)
d16	-		-		-		-	
d17	-		-		-		-	
d18	-		-		-		-	
d19	-		-		-		-	
d20	-		-		-		-	
d21	[1] 1024.0		[1] 1023.5		[1] 6443.5		[1] 6435.3	
d22	-		-		-		-	
d23	-		-		-		-	
d24	-		-		-		-	
d25	143.3	(67.1)	142.8	(67.1)	892.3	(418.6)	888.3	(418.7)

Figura 5.16: (Experiment 15) Valor estimat de les mitjanes i desviacions d'y i z, per a CD. (Per a cada estimador, mitjana i desviació de les 5 rèpliques).

RS								
	$\hat{E}(y)$		$\widehat{Dev}(y)$		$\hat{E}(z)$		$\widehat{Dev}(z)$	
d1	1.0	(0.0)	0.0	(0.0)	7.3	(0.1)	4.0	(0.2)
d2	1.0	(0.0)	0.0	(0.0)	7.9	(0.2)	6.0	(0.5)
d3	1.0	(0.0)	0.0	(0.0)	13.0	(0.1)	10.2	(0.6)
d4	1.4	(0.0)	0.8	(0.0)	80.0	(1.7)	82.8	(2.1)
d5	1.8	(0.0)	1.2	(0.0)	112.3	(5.6)	120.9	(4.9)
d6	1.8	(0.0)	1.2	(0.0)	122.1	(1.6)	127.1	(2.2)
d7	1.1	(0.0)	0.4	(0.0)	48.2	(0.6)	44.7	(0.9)
d8	1.2	(0.0)	0.5	(0.0)	61.9	(1.4)	53.4	(2.0)
d9	2.8	(0.1)	2.2	(0.1)	233.9	(7.1)	227.2	(7.1)
d10	6.2	(0.4)	5.7	(0.4)	572.4	(37.5)	574.6	(37.7)
d11	1.0	(0.0)	0.1	(0.0)	22.9	(0.4)	17.4	(0.9)
d12	291.8	(50.8)	291.3	(50.8)	29441.2	(5128.6)	29431.1	(5131.3)
d13	5.1	(0.2)	4.5	(0.2)	453.7	(21.7)	458.6	(22.3)
d14	6.4	(0.1)	5.9	(0.1)	596.0	(13.5)	593.6	(13.1)
d15	1.1	(0.0)	0.3	(0.0)	41.1	(0.2)	38.1	(0.5)
d16	93.3	(14.9)	92.8	(14.9)	9381.1	(1507.9)	9372.3	(1505.3)
d17	38.6	(5.1)	38.1	(5.1)	3894.1	(512.9)	3863.1	(513.1)
d18	19.4	(1.7)	18.9	(1.7)	1906.8	(175.0)	1913.3	(175.4)
d19	[1] 2048.0		[1] 2047.5		[1] 206995.0		[1] 206988.5	
d20	8.8	(0.3)	8.3	(0.3)	840.3	(34.9)	835.8	(34.4)
d21	4.4	(0.3)	3.9	(0.3)	421.9	(32.6)	396.3	(32.7)
d22	-		-		-		-	
d23	35.6	(3.7)	35.1	(3.7)	3550.3	(372.2)	3547.5	(371.9)
d24	348.3	(210.1)	347.8	(210.1)	35200.7	(21245.0)	35172.8	(21241.9)
d25	1.4	(0.0)	0.7	(0.0)	79.9	(2.1)	75.2	(2.4)

Figura 5.17: (Experiment 15) Valor estimat de les mitjanes i desviacions d'y i z, per a RS.

CC								
	$\hat{E}(y)$		$\widehat{Dev}(y)$		$\hat{E}(z)$		$\widehat{Dev}(z)$	
d1	1.3	(0.0)	0.7	(0.0)	40.5	(0.7)	22.8	(0.9)
d2	1.3	(0.0)	0.6	(0.0)	41.3	(0.5)	20.3	(0.6)
d3	1.8	(0.0)	1.2	(0.0)	59.8	(0.4)	39.6	(0.5)
d4	12.0	(1.2)	11.5	(1.2)	407.8	(41.7)	392.1	(41.8)
d5	14.0	(1.1)	13.5	(1.1)	482.2	(38.2)	465.0	(38.2)
d6	22.1	(1.3)	21.6	(1.3)	737.2	(42.2)	723.8	(42.1)
d7	14.2	(1.2)	13.7	(1.2)	502.5	(43.6)	485.5	(43.6)
d8	107.9	(40.7)	107.4	(40.7)	3641.9	(1374.4)	3628.1	(1374.1)
d9	69.4	(11.6)	68.9	(11.6)	2424.9	(403.5)	2407.4	(403.7)
d10	100.8	(13.6)	100.3	(13.6)	3497.4	(472.7)	3479.5	(472.4)
d11	2.8	(0.1)	2.2	(0.1)	95.8	(2.6)	77.0	(2.7)
d12	[1] 2048.0		[1] 2047.5		[1] 70806.0		[1] 70788.3	
d13	49.0	(5.5)	48.5	(5.5)	1630.3	(183.0)	1616.6	(183.2)
d14	270.4	(154.9)	269.9	(154.9)	9078.2	(5200.2)	9064.1	(5200.2)
d15	5.5	(0.1)	5.0	(0.1)	193.0	(4.8)	174.1	(4.8)
d16	[1] 1024.0		[1] 1023.5		[1] 34308.0		[1] 34294.8	
d17	[1] 2048.0		[1] 2047.5		[1] 73872.0		[1] 73851.1	
d18	89.7	(25.4)	89.2	(25.4)	2988.4	(848.1)	2974.2	(848.2)
d19	-		-		-		-	
d20	764.6	(261.7)	764.1	(261.7)	25615.7	(8779.6)	25602.4	(8779.6)
d21	[4] 1194.7	(591.2)	[4] 1194.2	(591.2)	[4] 43322.6	(21372.5)	[4] 43303.7	(21373.2)
d22	-		-		-		-	
d23	1262.9	(740.0)	1262.4	(740.0)	42187.6	(24720.7)	42173.0	(24719.3)
d24	-		-		-		-	
d25	12.5	(0.9)	12.0	(0.9)	443.5	(32.8)	424.7	(32.8)

Figura 5.18: (Experiment 15) Valor estimat de les mitjanes i desviacions d'y i z, per a CC.

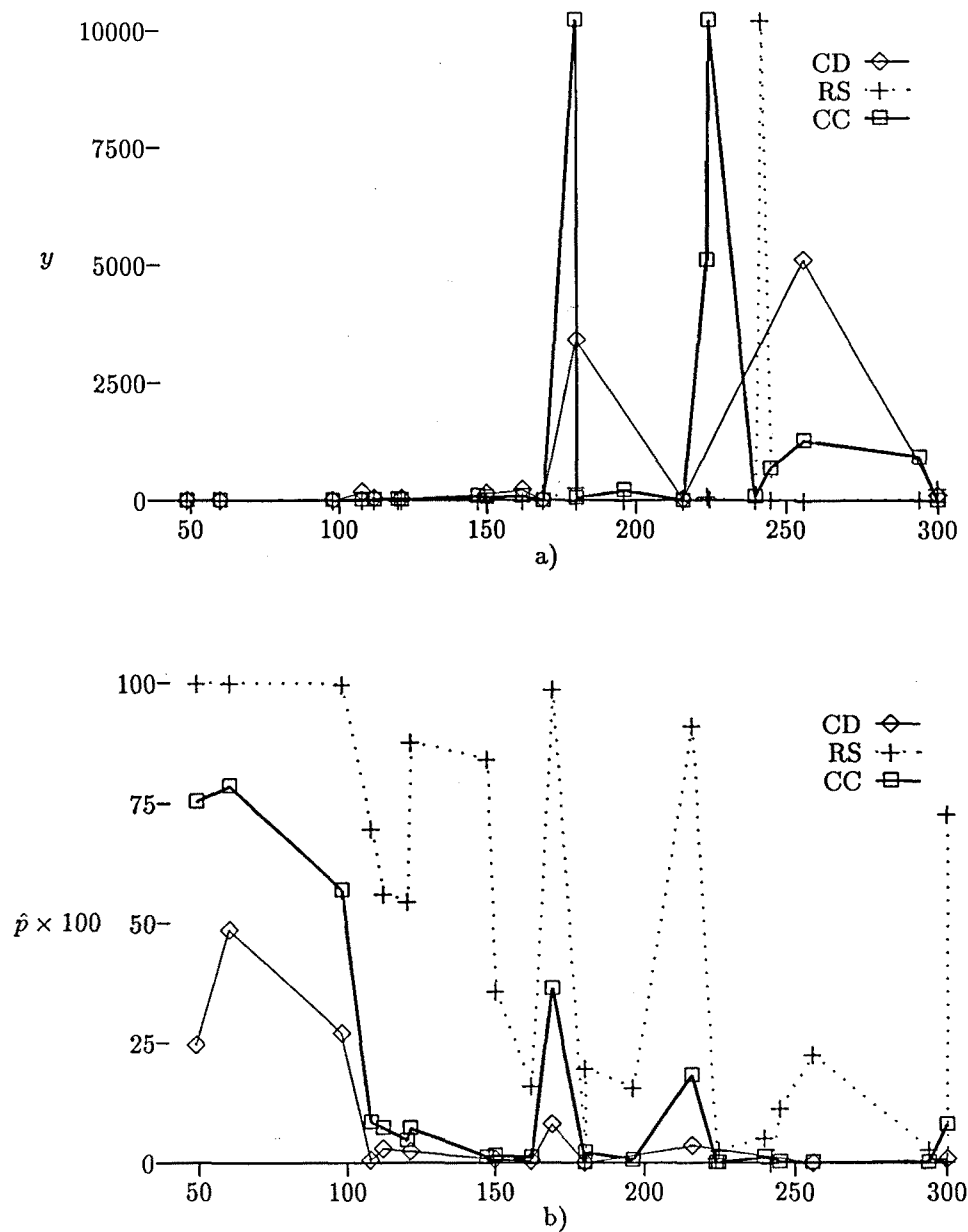


Figura 5.19: (Experiment 15) Per a les tres estratègies, evolució amb el tamany  $vb$  del problema a) del nombre d'intents esperat fins la primera solució  $\hat{E}(y)$ , i b) de la probabilitat  $\hat{p}$  estimada d'èxit.

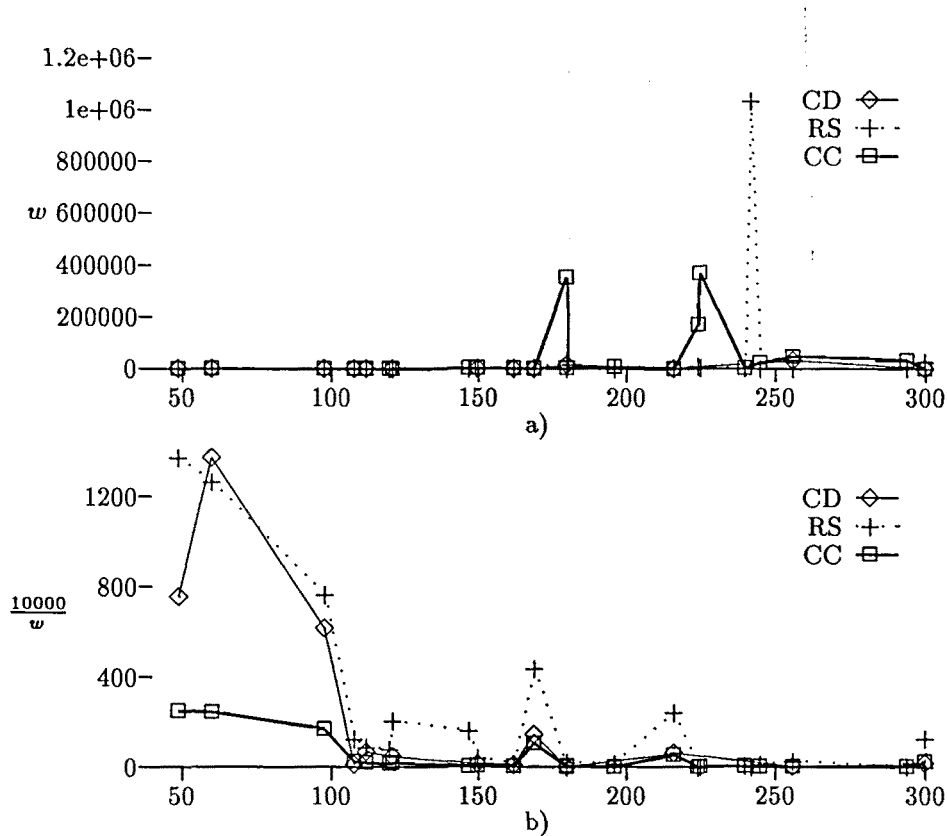


Figura 5.20: (Experiment 15) Per a les tres estratègies, evolució amb el tamany  $vb$  del problema del valor estimat a) del cost esperat  $E(z)$  fins la primera solució, i b) del seu estimador invers  $10000/w$ . c) Ampliació de l'anterior.



	$vb$	$\hat{E}(z)$
d26	300	8213.8
d28	320	7907.0
d29	324	451.4
d30	324	206807.0
d31	336	-
d32	338	69080.7
d33	338	-
d34	338	1923.8
d35	343	10291.8
d36	360	-
d37	360	17170.2
d38	361	-
d39	363	-
d40	364	-
d41	384	-
d42	392	25814.4
d43	396	206975.0
d44	420	34409.3
d45	432	68942.3
d46	441	85.5
d47	441	206799.0
d48	448	68902.0
d49	448	-
d50	450	-
d51	450	-
d52	480	-
d53	480	206783.0
d54	484	-
d56	486	-
d57	490	-

Figura 5.21: (Experiment 17) Valor estimat de l'esperança de  $z$  per als problemes amb  $vb \leq 500$ , amb l'algorisme guanyador ( $F_{uq}$  i  $RS^{std}$ ) (1 rèplica).

	<i>vb</i>	Resolt?		<i>vb</i>	Resolt?
d58	507	-	d94	726	-
d59	507	-	d95	728	-
d60	512	-	d96	729	-
d61	525	sí	d97	735	-
d62	525	-	d98	750	-
d63	528	sí	d99	750	sí
d64	528	-	d100	756	-
d65	529	-	d101	760	-
d66	539	-	d102	768	-
d67	540	-	d103	768	-
d68	540	sí	d104	792	-
d69	560	-	d105	792	-
d70	578	-	d106	810	-
d72	588	sí	d108	845	-
d73	600	-	d109	847	-
d74	600	-	d110	864	-
d75	605	sí	d111	882	-
d76	605	-	d112	882	-
d77	605	-	d113	882	-
d78	612	-	d114	896	-
d79	625	-	d115	900	-
d80	630	-	d116	900	-
d81	630	-	d117	900	-
d82	637	-	d118	918	-
d83	640	-	d119	924	-
d84	640	-	d120	945	-
d85	648	-	d121	960	-
d86	648	-	d122	960	-
d87	675	-	d123	961	sí
d88	676	sí	d124	961	-
d89	676	-	d125	961	-
d90	676	-	d126	968	-
d91	720	-	d127	968	-
d92	722	-	d128	972	-
d93	726	-	d129	1000	-

Figura 5.22: (Experiment 18) Problemes amb  $vb \leq 1000$  resolts per l'algorisme guanyador.

## Capítol 6

# Dissenys màximament equilibrats

En els capítols anteriors s'ha fet una avaluació de l'ús de l'optimització combinatòria per a la generació de dissenys en blocs incomplets equilibrats (DBs). Des d'aquest punt de vista, donat un conjunt de valors per als paràmetres  $(v, b, u)$ , era possible definir una funció de cost  $F$  tal que el seus òptims en l'espai de totes les configuracions  $A-(v \times b)$  fossin dissenys  $DB(v, b, u)$ . El rang de paràmetres admissibles, però, quedava restringit per la necessitat d'aconseguir files, columnes i correlacions *estrictament* uniformes.

Si acceptem que aquestes propietats es satisfacin amb el criteri de *màxima distribució*, el nou tipus de configuració que obtenim, els *dissenys màximament equilibrats* (DMEs), constitueix un super-conjunt dels dissenys en blocs que admet fins a dos valors consecutius en els seus descriptors. D'aquesta manera, *qualsevol* conjunt  $(v, b, u)$  de paràmetres és, *a priori*, admissible.

En aquest capítol es defineix el nou tipus de dissenys, i s'analitzen algunes de les seves propietats. Amb les mateixes mesures del capítol 2, i revisant les condicions de validesa, obtenim un repertori de funcions de cost adequades a la nova definició. Finalment, s'aplica la millor estratègia de cerca del capítol anterior per a la seva generació.

### 6.1 Definició i antecedents

La teoria combinatòria descriu diversos tipus de *pseudo*-dissenys per tal de cobrir aquells cassos en que els paràmetres  $(v, b, u)$  corresponents a un experiment particular no són admissibles (vegeu, p.ex., [Street & Street, 87]). Els més importants són els *dissenys equilibrats per parelles*,  $DEP(v, b, \lambda)$ , que forcen l'equilibri estricte (amb correlació  $\lambda$ ), però no posen restriccions a les files i columnes, i els *dissenys en blocs incomplets parcialment equilibrats*,  $DBIPE(i)-(v, b, r, k, \lambda_1, \dots, \lambda_i)$ , que imposen files i columnes estrictament uniformes (amb  $r$  i  $k$  uns, respectivament), però admeten fins a  $i$  valors diferents  $(\lambda_1, \dots, \lambda_i)$  a les correlacions entre parelles de files.

**Definició 6.1** *Sigui  $S$  un conjunt amb  $v$  elements, i  $B$  una família de  $b$  subconjunts o blocs del conjunt  $S$ . Per a valors fixos d' $r$ ,  $k$ , i  $\lambda$ , direm que  $B$  és un disseny màximament equilibrat si i només si*

- i) *cada element pertany a  $r$  o  $r + 1$  blocs,*
- ii) *cada bloc conté  $k$  o  $k + 1$  elements, i*
- iii) *qualsevol parella d'elements apareix en  $\lambda$  o  $\lambda + 1$  blocs.*

Comparant-la amb la definició 1.1, cal remarcar que s'han eliminat les restriccions sobre  $\lambda$  i  $k$ , i que la condició i) d'uniformitat de les files s'hi inclou explícitament ja que, com veurem, en aquest cas no es deriva de les altres dues.

Prement  $(v, b, u)$  com a paràmetres independents, amb  $0 \leq u \leq vb$ , els valors  $r$ ,  $k$  i  $\lambda$  de la definició venen lligats per relacions de consistència. Si  $u$  és el nombre d'uns de la matriu d'incidència, aleshores definim  $r = \lfloor u/v \rfloor$  i, amb  $v_0 = u \bmod v$ , hi haurà  $v_0$  files amb  $r + 1$  uns, i la resta en tindran  $r$ . Anàlogament, si definim  $k = \lfloor u/b \rfloor$ , aleshores el nombre de columnes amb  $k + 1$  uns és  $b_0 = u \bmod b$ , i la resta en tindran  $k$ . Amb aquesta distribució (amb columnes uniformement distribuïdes), el nombre  $p_v^o$  de parelles verticals és mínim i val  $p_v^o = b \binom{k}{2} + b_0 k$ , d'on  $\lambda = \lfloor p_v^o / f \rfloor$ , i el nombre de files amb correlació  $\lambda + 1$  és  $f_0 = p_v^o \bmod f$ . La resta correlaran  $\lambda$ . En el cas dels DMEs, doncs, donat un conjunt de paràmetres, els *descriptors* del disseny són  $[u, (r, k, \lambda), (v_0, b_0, f_0)]$ , i en el cas particular dels DBs, els descriptors del darrer grup són tots zero.

La definició de DME, doncs, no imposa restriccions d'admissibilitat.

## 6.2 Algunes propietats dels dissenys màximament equilibrats

Com en el cas dels DBs, donat un conjunt  $(v, b, u)$  de paràmetres, l'existència d'un  $DME(v, b, u)$  no està garantida. Algunes de les propietats que es descriuen a continuació estan formulades com a condicions d'existència.

La primera és un exemple del tipus de situacions "patològiques" amb que ens podem trobar.

**Proposició 6.1** *Donat un conjunt  $(v, b, u)$  de paràmetres, si*

$$1 < \bar{u} < v - 1, \quad (6.1)$$

*aleshores no existeix cap  $DME(v, b, u)$ .*

En aquesta situació la matriu d'incidència és pràcticament plena. Per no violar la uniformitat de files i columnes, els  $\bar{u}$  zeros s'han de posar en files i columnes

diferents. Si ho fem així, però, quan es compleix la condició (6.1) les correlacions entre files prenen tres valors diferents. Un exemple d'aquest fet és el cas (6, 10, 58). A l'apartat 6.3, veurem un fenomen similar.

**Proposició 6.2** *Donat un conjunt  $(v, b, u)$  de paràmetres tal que  $\lambda > 0$ , si*

$$b_0 = 0, f_0 = 0, \quad i \quad v_0 \neq 0, \quad (6.2)$$

*aleshores no existeix cap  $DME(v, b, u)$ .*

Aquesta restricció es basa en l'equació (1.3), que expressa que si columnes  $i$  correlacions són *estrictament* uniformes, *aleshores* les files també ho han de ser. Un exemple d'aquesta circumstància és el cas (6,10,40), que té com a descriptors [40,(6,4,4),(4,0,0)]. O bé, amb dimensions més grans, el cas (15,21,126), amb descriptors [126,(8,6,3),(6,0,0)].

El següent resultat expressa una relació important amb l'espai de configuracions *dual*. Com veiem al capítol 2, el nombre de *quàdruples* d'una configuració  $A - v \times b$  amb  $o$  uns ve donat per

$$Q(A) = \sum_i \sum_{k \neq i} \binom{\lambda_{ik}}{2}, \quad (6.3)$$

amb  $\lambda_{ik}$  la correlació entre les files  $i$  i  $k$ . Si  $o = u$ , i els uns estan uniformement distribuïts entre columnes, i les parelles verticals estan uniformement distribuïdes entre parelles de files (equilibri), *aleshores*  $Q(A) = Q_e$ , amb

$$Q_e \equiv f \binom{\lambda}{2} + f_0 \lambda, \quad (6.4)$$

la *cota inferior* de  $Q$  per a una configuració equilibrada amb  $o = u$  uns (proposició 2.3).

Fent un raonament *dual* a l'anterior, les quàdruples d'una configuració també es poden interpretar com el nombre de parelles que es poden formar amb parelles *horitzontals* d'uns ("boles") entre parelles de *columnes* ("pots"). Des d'aquest punt de vista, si  $\mu_{jl}$  és la correlació entre les columnes  $j$  i  $l$ , el nombre de quàdruples de la configuració  $A$  és

$$Q(A) = \sum_j \sum_{l \neq j} \binom{\mu}{2}. \quad (6.5)$$

Direm que una configuració és *lligada* quan les correlacions entre parelles de columnes estan màximament distribuïdes. Amb  $g = \binom{b}{2}$  parelles de columnes, i  $p_h$  parelles horitzontals, la condició de màxima distribució queda descrita per  $\mu(p_h)$  i  $g_0(p_h)$ . Per a  $(g, p_h)$  tals que  $g_0(p_h) = 0$  totes les correlacions entre columnes valen  $\mu(p_h)$  i direm que  $A$  és *estrictament* lligada. Si  $o = u$ , i les files estan uniformement distribuïdes, i  $A$  és lligada, *aleshores*  $Q(A) = Q_u$ , amb

$$Q_u \equiv g \binom{\mu}{2} + g_0 \mu, \quad (6.6)$$

on  $[\mu, g_0]$  són els descriptors de *lligam* que es deriven dels paràmetres  $(v, b, u)$ . En conseqüència, per dualitat,  $Q_{ll}$  és la *cota inferior* de  $Q$  per a una configuració lligada amb  $o = u$  uns.

Així, si  $Q_e \neq Q_{ll}$ , totes les condicions esmentades no es poden satisfer *simultàniament*. Donat que els dissenys es defineixen en termes d'equilibri, tenim

**Teorema 6.3** *Es condició necessària per a l'existència d'un disseny màximament equilibrat amb paràmetres  $(v, b, u)$  que les expressions  $Q_e$  (eq. 6.4) i  $Q_{ll}$  (eq. 6.6) satisfacin*

$$Q_e \geq Q_{ll}. \quad (6.7)$$

Per exemple, com a conseqüència d'aquesta propietat, podem determinar que els dissenys amb paràmetres  $(6, 10, 40)$ ,  $(6, 10, 50)$  i  $(6, 10, 51)$  no existeixen. El primer d'ells ja estat mencionat anteriorment com exemple de la proposició 6.2<sup>1</sup>. En el cas dels DBs, un fet curiós a remarcar és que els 28 primers casos coneguts de no existència llistats a [Mathon & Rosa, 90], satisfan la condició (6.7) *amb igualtat*.

Els enunciats que es formulen a continuació són conseqüència immediata del teorema que acabem de formular:

- El *dual* d'un disseny equilibrat és una configuració lligada.
- Si existeix un DME amb  $Q_e > Q_{ll}$ , el seu dual *no* és un DME.
- Si existeix un  $DME(v, b, u)$  amb  $Q_e = Q_{ll}$ , aleshores es tracta d'un DME lligat, i el seu dual també és un DME, amb paràmetres  $(b, v, u)$ .

I es verifica també que

- Si un DME és quadrat ( $v = b$ ), aleshores  $Q_e = Q_{ll}$ . Però  $Q_e = Q_{ll}$  *no* implica  $v = b$ .

Dels 25 primers DBs de la taula 4.1, només tres (d2, d4 i d9) satisfan la igualtat anterior amb  $v \neq b$ . En el cas dels DBs, donat que no existeix cap disseny amb  $v > b$  (pel teorema de Fisher), el dual d'un DB *lligat* només és un DB si  $v = b$ , però sempre és un DME.

Així, com a condició d'existència, el teorema 6.3 és totalment inefectiu quan  $v = b$ , i com s'il·lustra més endavant sobre els espais de dimensions  $(6, 10)/(10, 6)$ , resulta molt poc restrictiu per a  $v < b$ , però fortament restrictiu per a  $v > b$ . La condició de Fisher, doncs, sembla manifestar-se ara de manera relaxada.

En termes de DMEs, les equacions (2.64, 2.65) de l'apartat 2.3.1, es poden reescriure com segueix.

**Teorema 6.4** *Donat un conjunt de paràmetres  $(v, b, u)$  admissible (amb  $v_0 = b_0 = f_0 = 0$ ), l'existència d'un disseny en blocs  $DB(v, b, u)$  és condició necessària i suficient per a l'existència dels dissenys màximament equilibrats  $DME(v, b, u - 1)$  i  $DME(v, b, u - 1)$ .*

<sup>1</sup>Les dues propietats, però, no sempre coincideixen. En l'altra cas mencionat anteriorment, per exemple, amb paràmetres  $(15, 21, 126)$ , es verifica que  $Q_e = 315$  mentre que  $Q_{ll} = 306$ .

Donat el seu abast, aquest teorema, més que establir una condició d'existència per als DBEs, descriu una propietat del veïnatge dels DBs. La seva demostració s'obté verificant que una configuració màximament distribuïda (és a dir, un DME), amb  $u \pm 1$  uns, és efectivament veïna d'un  $\text{DB}(v, b, u)$ . Fent una anàlisi similar a la de la figura 2.5, es verifica que l'increment local de l' $u$  que sobra/falta és exactament igual a l'increment local òptim corresponent en el DB. A més, per la propietat transitiva, l'existència d'un d'aquests dos DBEs implica l'existència de l'altra.

La propietat anterior no és extensiva als DMEs en el cas general ja que, com es mostra a continuació, els increments locals òptims d'un DME poden pendre més de dos valors diferents. Enllaçant amb l'apartat 2.3, ens fixem novament en les quàdruples, que és el cas més complicat (una situació similar es produeix per a la resta de mesures). Per a una configuració  $A$  donada els increments locals es poden expressar com

$$\Delta^{ij}Q(A)|_{x_{ij}=1} = \sum_{\{l:l \neq i, x_{lj}=1\}} (\lambda_{il} - 1) \quad (6.8)$$

$$\Delta^{ij}Q(A)|_{x_{ij}=0} = \sum_{\{l:l \neq i, x_{lj}=1\}} \lambda_{il}, \quad (6.9)$$

on els sumatoris s'extenen sobre els  $k_j - x_{ij}$  uns de la columna  $j$ . Sota la hipòtesi d'optimalitat, redefinida en termes de DMEs, els descriptors de columnes i correlacions poden pendre ara dos valors i, depenent de la posició  $(i, j)$ , els increments locals òptims queden compresos en els rangs següents:

$$\Delta^{ij}Q^*|_{x_{ij}=1} \in [(k-1)(\lambda-1), k\lambda] \quad (6.10)$$

$$\Delta^{ij}Q^*|_{x_{ij}=0} \in [k\lambda, (k+1)(\lambda+1)]. \quad (6.11)$$

Si  $b_0$  o  $f_0$  valen 0, els rangs es fan més petits, i per a  $b_0 = f_0 = 0$ , es redueixen a un sol valor:  $\Delta^1Q^*$  quan  $x_{ij} = 1$ , i  $\Delta^0Q^*$  en cas contrari. Per la proposició 6.2, aleshores, quan  $b_0 = f_0 = 0$  només pot existir el disseny si  $v_0$  també val 0, de manera que el teorema 6.4 només aplica als DBs.

La darrera propietat que s'enuncia, que no implica condicions d'existència, presenta una casuística similar.

**Teorema 6.5** *Donat un  $\text{DME}(v, b, u)$ , la configuració complementària és un DME amb paràmetres  $(v, b, vb - u)$  si i només si*

$$v_0 = 0 \text{ o } (f_0 = 0 \text{ i } (v_0 = 1 \text{ o } v_0 = v - 1)). \quad (6.12)$$

Per a una configuració  $A$ , donades dues files  $i$  i  $k$  qualsevols que correlin  $\lambda_{ik}$  i tinguin  $r_i$  i  $r_k$  uns, respectivament, la correlació  $\bar{\lambda}_{ik}$  a la configuració complementària  $\bar{A}$  ve donada per (eq. 2.58)

$$\bar{\lambda}_{ik} = b + \lambda_{ik} - r_i - r_k. \quad (6.13)$$

Si  $A$  és un DME, quan  $v_0 \neq 0$  algunes files tenen  $r$  uns i unes altres en tenen  $r + 1$ , i quan  $f_0 \neq 0$ , passa el mateix amb les correlacions ( $\lambda$  i  $\lambda + 1$ ). Aplicant

l'expressió anterior a *totes* les parelles de files  $i$ , analitzant tots els casos, s'obté la condició (6.12).

Així, en termes generals, quan per a un conjunt de paràmetres  $(v, b, u)$  no existeix cap disseny, per al conjunt de paràmetres complementari  $(v, b, vb - u)$  en pot existir algun, i el fet de que la *configuració* complementària d'un  $\text{DME}(v, b, u)$  no sigui un DME, no vol dir que no existeixi cap  $\text{DME}(v, b, vb - u)$ . Tornant al cas  $(6, 10, 40)$ , per exemple, tot i que sabem que no existeix, és possible construir el complementari  $\text{DME}(6, 10, 20)$ , i els dissenys  $\text{DME}(6, 10, 27)$  i  $\text{DME}(6, 10, 33)$ , amb paràmetres complementaris però que no satisfan (6.12), existeixen independentement. El complementari d'un d'ells, però, no és solució de l'altre. L'espai  $(6, 10)/(10, 6)$  s'avalua exhaustivament a l'apartat 6.4, en termes també d'aquesta propietat.

### 6.3 Funcions de cost per a la generació de dissenys màximament equilibrats

Les restriccions que hem d'imposar a una funció de cost per a la generació de DMEs depèn dels descriptors  $(v_0, b_0, f_0)$ . L'objectiu d'aquest apartat, però, és formular una expressió general que, tenint en compte el "pitjor cas", serveixi per a tota la resta. Dissenyarem, doncs, una funció de cost per al cas en que tots tres són diferents de 0.

Enllaçant amb l'apartat 2.4.2, una mesura  $F(\rho_u, \rho_t, \rho_h, \rho_v, \rho_q, \rho_{\bar{q}})$  és una funció de cost per a la generació de dissenys  $\text{DME}(v, b, u)$  si els termes que la componen garanteixen l'optimalitat, i té un mínim puntual a  $o = u$  (la minimalitat global, com abans, queda garantida per la convexitat dels increments locals). Per als DMEs, doncs, i sobre totes les configuracions amb  $o = u$  uns, la minimització de  $Q$  garanteix l'equilibri, però, donat que ara  $\lambda$  pot valdre 0, la uniformitat de columnes s'ha de forçar explícitament amb el terme  $P_v$  (proposició 2.3). A més, com s'il·lustra a la figura 6.1,... en el cas dels DMEs equilibri i columnes uniformes ja no garanteixen la uniformitat de les files, de manera que també cal incloure el terme  $P_h$ . Així, en el cas dels DMEs, el teorema 2.4 es tradueix a

**Teorema 6.6** *Donat un conjunt de paràmetres  $(v, b, u)$ , amb  $0 \leq u \leq vb$ , una configuració  $A \in \mathcal{A} - v \times b$  és un disseny màximament equilibrat  $\text{DME}(v, b, u)$  si i només si*

$$U(A) = u, \quad (6.14)$$

$$P_h(A) = P_h^*, \quad (6.15)$$

$$P_v(A) = P_v^* \quad i \quad (6.16)$$

$$Q(A) = Q^*. \quad (6.17)$$

A més. el corol·lari 2.4.1 ja no aplica donat que, com hem vist,

- Una configuració amb  $o = u$  uns pot tenir correlacions uniformes sense ser-ho les files ni les columnes.



A1, DME(6,10,27), [27, (4,2,1)(3,7,9)]													A2, files no uniformes															
1			1	1		1		4	2	1	1	1	2			1		1	1	1		1	5	2	2	2	1	2
1		1					1	1	4	1	1	2	2			1	1	1		1			4	2	1	1	1	
1	1	1		1		1		5	2	2	2					1		1	1	1			4	1	1	2		
		1	1		1	1		4	1	2						1		1	1			1	4	2	2			
1			1	1		1		1	5	2						1	1					1	1	4	2			
			1				1	1	1	1	1	5				1	1		1	1	1	1	6					
2 3 3 3 2 2 3 3 3 3 27													3 2 3 3 3 3 3 2 2 3 27															

Figura 6.1: Exemples de configuracions-(6,10). La primera correspon a un DME amb els paràmetres i descriptors que es mostren. La segona, malgat que té columnes i correlacions uniformes, enlloc de tenir 3 files amb 5 uns i 3 files amb 4 uns, té fins a 3 valors diferents d' $r_i$ .

- La configuració complementària d'un DME pot no ser un DME.
- Els increments locals òptims en un DME poden pendre diversos valors.

Una conseqüència important del segon punt és que el terme  $\bar{Q}$  no pot aparèixer a la funció de cost. I una conseqüència important del tercer punt és que les funcions de tipus  $H$  de l'apartat 2.6 ja no es poden aplicar.

La taula 6.1 mostra els valors i els increments òptims per als DMEs. Ara, en relació als increments locals, es verifica tan sols que

$$\Delta^{ij}M^*|_{x_{ij}=1} \leq \Delta^-M^* \quad (6.18)$$

$$\Delta^{ij}M^*|_{x_{ij}=0} \geq \Delta^+M^*. \quad (6.19)$$

Pel que fa al mínim puntual, n'hi ha prou amb que un sol terme  $M$  de la funció de cost compleixi  $\Delta(M|U)^{min}(u-1) < \Delta(M|U)^{min}(u)$ , ja que la resta, en el pitjor dels casos, són constants. Però, amb  $v_0$ ,  $b_0$  i  $f_0$  diferents de 0, no es satisfia cap de les condicions de multiplicitat, i ens cal recórrer necessàriament a  $\Delta(P_t|U)(o)$  que és estrictament creixent. Així, el teorema 2.5 s'ha de substituir pel següent:

**Teorema 6.7 Donat**

un conjunt de paràmetres  $(v, b, u)$ , la mesura  $F(\rho_u, \rho_t, \rho_h, \rho_v, \rho_q, \rho_{\bar{q}})$  és una funció de cost per a la generació de dissenys màximament equilibrats si

$$\rho_t, \rho_h, \rho_v, \rho_q > 0 \quad (6.20)$$

$$\rho_{\bar{q}} = 0, \quad (6.21)$$

$i$

$$\Delta^-F^* = \Delta(F|U)^m(u-1) < 0 \quad (6.22)$$

$$\Delta^+F^* = \Delta(F|U)^{min}(u) > 0. \quad (6.23)$$

$M$	$M^*$	$\Delta^- M^*$	$\Delta^+ M^*$
$U$	$u$	1	1
$P_t$	$\binom{u}{2}$	$u - 1$	$u$
$P_h$	$v \binom{r}{2} + v_0 r$	$\begin{cases} r - 1 & \text{si } v_0 = 0 \\ r & \text{si } v_0 \neq 0 \end{cases}$	$r$
$P_v$	$b \binom{k}{2} + b_0 k$	$\begin{cases} k - 1 & \text{si } b_0 = 0 \\ k & \text{si } b_0 \neq 0 \end{cases}$	$k$
$Q$	$f \binom{\lambda}{2} + f_0 \lambda$	$\begin{cases} (k - 1)(\lambda - 1) + f_0 & \text{si } b_0 = 0, f_0 < k - 1 \\ (k - 1)\lambda & \text{si } b_0 = 0, f_0 \geq k - 1 \\ k(\lambda - 1) + f_0 & \text{si } b_0 \neq 0, f_0 < k \\ k\lambda & \text{si } b_0 \neq 0, f_0 \geq k \end{cases}$	$\begin{cases} k\lambda & \text{si } f_0 \leq f - k \\ k(\lambda + 1) - (f - f_0) & \text{si } f_0 > f - k \end{cases}$

Taula 6.1: Valors i increments òptims de les mesures  $U$ ,  $P_t$ ,  $P_h$ ,  $P_v$  i  $Q$ , en termes dels descriptors d'un DME( $v, b, u$ ).

El seu valor òptim és

$$\begin{aligned} F^* &= (F|U)^{min}(u) = \\ &= \rho_u U^* + \rho_t P_t^* + \rho_h P_h^* + \rho_v P_v^* + \rho_q Q^*, \end{aligned} \tag{6.24}$$

amb els valors òptims donats per la taula 6.1.

L'única estructura admesa és, en conseqüència,  $F_{uthvq}$ , i la seva composició es fixarà experimentalment.

### 6.4 Generació experimental de dissenys màximament equilibrats

A l'hora d'escollir el repertori de prova, i quan no es digui altrament, els parametres d'un DME s'han agafat a l'atzar de manera que  $vb \leq 300$  i  $u \in [b + 1, vb - b]$ , prescindint així dels casos més trivials.

El primer experiment que presentem correspon a la selecció dels coeficients de composició d' $F_{uthvq}$ . Dels 7 DMEs que fem servir d'entrenament (e0,...,e6), els paràmetres de (e1,e2,e5 i e6) s'han escollit a l'atzar, e0 és complementri d'e1, e3 és quadrat, i e4 és dual d'e5.

#### Experiment 19

**Objectius:** Cerca d'una bona composició per a  $F_{uthvq}$ .

**Problemes:** Conjunt de DMEs d'entrenament (e0,...,e6).

**Funcions:**  $F_{uthvq}$ .

**Estratègies:** CL.

**Rèpliques:** 1.

**Comentaris:** Realitzat conjuntament amb l'experiment 5 del capítol 5.

Els resultats que s'obtenen, que no es mostren en detall, són molt poc sensibles a la composició. De fet, per a aquesta funció, l'elecció de coeficients s'ha fet conjuntament amb els DBs de prova, i cap de les decisions que s'han pres en aquell cas (apartat 5.2.1) ha vingut afectada per la presència dels DMEs. Per a les 30 composicions provades, els 4 casos que es resolen (e1, e2, e4 i e5) ho fan sempre amb un cost estimat molt baix. Mentre els 3 casos restants fracassen sempre. Els coeficients que s'obtenen són doncs, com aleshores,  $(\alpha_u, 1, 9, 3, 7, 0)$ .

En el proper experiment s'analitza experimentalment la casuística de l'apartat 6.2, per a tots els problemes de l'espai  $\mathcal{A} - 6 \times 10$  i el seu dual  $\mathcal{A}' - 10 \times 6$ , amb la millor estratègia obtinguda en el capítol 5: RS amb paràmetres estàndard.

### Experiment 20

**Objectius:** Anàlisi de la casuística dels DMEs.

**Problemes:** Tots els conjunts de paràmetres amb  $(v, b) = (6, 10)$  i  $(10, 6)$ .

**Funcions:**  $F_{uthvq}$ .

**Estratègies:** RS<sup>std</sup>.

**Rèpliques:** 1.

La taula 6.2 sintetitza els resultats. Amb l'excepció de  $(6, 10, 41)$  i  $(6, 10, 52)$ , tots els casos no resolts es justifiquen per alguna condició de no existència. El primer,  $(6, 10, 41)$  no ha estat exhaustivament verificat. Però el segon,  $(6, 10, 52)$ , il·lustra una situació similar a la proposició 6.1, com es mostra a la figura 6.2. A banda d'isomorfismes, la distribució dels zeros no presenta alternatives. Però les correlacions dels uns no estan equilibrades.

Tot i que el tamany del problema és el mateix, les diferències de cost són sorprenentment grans. Però no s'ha pogut trobar cap dependència evident que relacioni la dificultat del problema amb els descriptors del disseny. Fins i tot algunes configuracions amb poc uns, trivials de construir a mà, són difícils per a la cerca. La proposició 6.1, feta a mida pels casos que la verifiquen, no aporta més informació. La proposició 6.2 és més útil, ja que tot i correspondre a una circumstància molt concreta, els exemples que filtra són difícils de verificar altrament. El teorema 6.3, com hem dit, discrimina molt a l'espai dual  $\mathcal{A}'$ , però només per a valors de  $u$  tals que  $\lambda > 0$ . El teorema 6.4 queda exemplificat pels casos  $(6, 10, 30)$  i  $(6, 10, 30 \pm 1)$ . Tots tres existeixen, i s'han resolt. Finalment, la condició de complementarietat (teorema 6.5) es satisfà poques vegades.

El darrer experiment, aplica l'algorisme de cerca a un conjunt de DMEs de prova. Com hem fet amb els DBs grans, la cerca paral·lela s'atura tan bon punt un dels processadors troba un disseny.

		$\mathcal{A} - v \times b$			$\mathcal{A}' - b \times v$		
		DME( $v, b, u$ )		Compl?	DME( $v, b, u$ )		Compl?
$u$	$\bar{u}$	$\hat{E}(z)$	Exist?		$\hat{E}(z)$	Exist?	
0	60	1.0			1.0	1.0	
1	59	1.0			1.0	1.0	
2	58	2.0		no	1.9	- no [1]	no
3	57	2.0		no	2.0	- no [1]	no
4	56	2.0		no	2.0	- no [1,3]	no
5	55	2.0			2.0	- no [1,3]	no
6	54	24.6			2.7	- no [1,3]	no
7	53	291.5			2.7	- no [1,3]	no
8	52	93.3		no	2.2	- no [1]	no
9	51	39.1		no	2.6	2.4	no
10	50	3.0		no	8.3	4.1	
11	49	4.4		no	133.3	- no [3]	no
12	48	9.5			15.6	- no [3]	no
13	47	13.9		no	3.1	- no [3]	no
14	46	6.5		no	2.7	- no [3]	no
15	45	3.7		no	2.7	- no [3]	no
16	44	2.6		no	2.6	- no [3]	no
17	43	2.7		no	2.4	- no [3]	no
18	42	16.7			2.5	- no [3]	no
19	41	171.5		no	30.0	- no [3]	no
20	40	6.3		no	8.6	246.8	
21	39	66.2		no	83.4	2487.5	no
22	38	27.1		no	133.8	- no [3]	no
23	37	38.7		no	- no [3]	- no [3]	no
24	36	10.8			- no [3]	- no [2,3]	no
25	35	22.3		no	- no [3]	- no [3]	no
26	34	21.9		no	- no [3]	- no [3]	no
27	33	18.3		no	- no [3]	- no [3]	no
28	32	23.5		no	- no [3]	- no [3]	no
29	31	595.6		no	- no [3]	- no [3]	no
30		24.7			4060.6		

Taula 6.2: (Experiment 20) Exploració de l'espai  $\mathcal{A} - 6 \times 10$  (esquerra), i dual (dreta). Entre  $u$  i  $\bar{u}$  es cobreix tot el repertori. En cada cas, resultat de la cerca (1 rèplica) i existència. Si no existeix, s'indica entre claus [] la proposició que ho estableix (per ordre d'aparició en el text). Independentment d'“Exist”, l'optimalitat del complementari, “Compl?”, es determina pel teorema 6.5

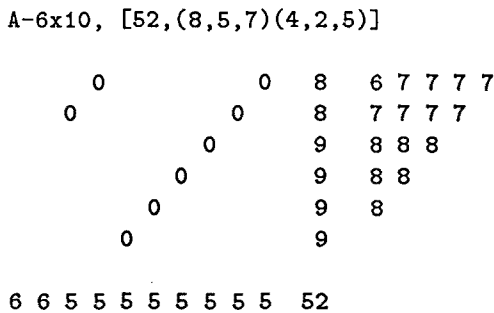


Figura 6.2: Amb paràmetres (6,10,52), distribució òptima dels zeros (totes les alternatives són isomorfes d'aquesta). Només d'aquesta manera es preserva la uniformitat dels uns entre files i columnes, però les correlacions no són equilibrades. El DME, doncs, no existeix.

Experiment 21

- Objectius: Generació de DMEs.
- Problemes: DMEs d'entrenament i prova (42 problemes).
- Funcions:  $F_{uthvq}$ .
- Estratègies:  $RS^{std}$ .
- Rèpliques: 1.
- Descripció: Experiment paral·lel truncat a la primera solució (no cal acabar la resta de descensos). Si no en trobem cap, tenim un fracàs total (cap disseny en  $N = 2048$  intents).

Paràmetres i resultats es llisten a la taula 6.3. El primer bloc de problemes correspon al conjunt d'entrenament, que s'ha mencionat abans. Pel que fa als problemes de prova, els 24 casos que segueixen (e7,...,e30) tenen les dimensions dels dissenys del capítol 5, amb uns escollits a l'atzar (el cas (6,10), naturalment, no es repeteix). En total 5 fracassos, un d'ells, e23, justificat. En el procés de selecció d'aquests problemes, els paràmetres que conduïen a  $Q_e < Q_{ll}$  eren rebutjats, en virtut del teorema 6.3, i es repetia la tria. El següent grup de problemes, però, (e31,...,e37) amb  $v > b$ , recullen els casos rebutjats des de l'espai dual. La proporció de casos irresolts en aquest grup (4 de 7), sembla indicar que el teorema de Fisher també es pot manifesta en els resultats experimentals: els DMEs amb  $v > b$  també són més difícils de resoldre. Finalment, l'últim bloc de problemes (e38,...,e41) correspon a configuracions veïnes dels dos DBs més difícils del capítol anterior, d19 i d22. Malhauradament (però com era d'esperar) no s'han pogut resoldre.

Després de tot, i en circumstàncies similars, els DMEs no semblen més difícils de generar que els DBs, especialment tenint en compte que no es pot utilitzar la funció  $F_{uq}$ , que era la que millors resultats donava en el capítol 5. Finalment, en el cas dels

	<i>vb</i>	<i>v</i>	<i>b</i>	<i>u</i>	<i>r</i>	<i>k</i>	$\lambda$	<i>v</i> <sub>0</sub>	<i>b</i> <sub>0</sub>	<i>f</i> <sub>0</sub>	<i>Q<sub>e</sub></i>	<i>Q<sub>II</sub></i>	Result?
e0	77	7	11	44	6	4	3	2	0	3	72	69	-
e1	77	7	11	33	4	3	1	5	0	12	12	7	sí
e2	90	6	15	35	5	2	1	5	5	10	10	0	sí
e3	100	10	10	85	8	8	7	5	5	5	980	980	-
e4	182	14	13	23	1	1	0	9	10	10	0	0	sí
e5	182	13	14	23	1	1	0	10	9	9	0	0	sí
e6	240	15	16	133	8	8	4	13	5	68	902	896	-
e7	49	7	7	39	5	5	4	4	4	6	150	150	sí
e8	98	7	14	55	7	3	3	6	13	18	117	105	sí
e9	108	9	12	89	9	7	7	8	5	35	1001	990	-
e10	112	8	14	49	6	3	2	1	7	7	42	35	sí
e11	120	6	20	74	12	3	6	2	14	12	297	270	sí
e12	121	11	11	63	5	5	2	8	8	40	135	135	sí
e13	147	7	21	118	16	5	13	6	13	2	1664	1644	-
e14	150	10	15	65	6	4	2	5	5	20	85	75	sí
e15	162	9	18	38	4	2	0	2	2	22	0	0	sí
e16	169	13	13	91	7	7	3	0	0	39	351	351	sí
e17	180	10	18	111	11	6	6	1	3	18	783	765	sí
e18	180	6	30	78	13	2	4	0	18	6	114	33	sí
e19	196	7	28	65	9	2	2	2	9	4	29	0	sí
e20	216	9	24	98	10	4	4	8	2	8	248	209	sí
e21	224	8	28	52	6	1	0	4	24	24	0	0	sí
e22	225	15	15	187	12	12	10	7	7	24	4965	4965	-
e23	240	6	40	160	26	4	16	4	0	0	1800	1768	- [2]
e24	242	11	22	100	9	4	3	1	12	15	210	174	sí
e25	245	7	35	193	27	5	20	4	18	20	4390	4310	sí
e26	256	16	16	197	12	12	9	5	5	36	4644	4644	-
e27	264	12	22	240	20	10	18	0	20	2	10134	10125	sí
e28	294	7	42	204	29	4	18	1	36	18	3537	3447	sí
e29	300	10	30	163	16	5	8	3	13	5	1300	1191	sí
e30	300	10	30	134	13	4	5	4	14	11	505	397	sí
e31	120	20	6	100	5	16	4	0	4	24	1236	1235	sí
e32	150	15	10	105	7	10	4	0	5	80	950	945	-
e33	180	18	10	143	7	14	6	17	3	34	2499	2497	-
e34	224	28	8	196	7	24	6	0	4	36	5886	5880	sí
e35	240	40	6	200	5	33	4	0	2	114	5136	5135	sí
e36	264	22	12	197	8	16	6	21	5	134	4269	4268	-
e37	182	14	13	155	11	11	9	1	12	28	3528	3520	-
e38	242	11	22	109	9	4	3	10	21	51	318	279	-
e39	242	11	22	111	10	5	4	1	1	5	350	317	-
e40	264	12	22	131	10	5	4	11	21	61	640	607	-
e41	264	12	22	133	11	6	5	1	1	6	690	649	-

Taula 6.3: (Experiment 21) Paràmetres, descriptors i resultats per als DMEs d’entrenament (bloc superior) i prova. Les darreres columnes mostren les cotes de les quàdruples i el resultat de la cerca. El cas e23 no existeix, per la proposició 6.2.

DMEs, donat que no coneixem més condicions d'existència, tots els casos no resolts corresponen a problemes oberts.

# Capítol 7

## Conclusions

En aquest treball hem presentat la proposta i comparació d'alguns models de xarxes neuronals optimitzadores per a la generació de dissenys en blocs, així com la seva extensió als dissenys màximament equilibrats. En aquest darrer capítol es descriuen les contribucions del treball i els resultats de l'avaluació experimental, seguides d'una discussió que permet fixar les línies futures de recerca.

### 7.1 Aportacions i altres treballs relacionats

Les aportacions principals d'aquest treball són les següents:

- L'ús de xarxes neuronals d'ordre superior a per a l'optimització combinatoria [Bofill & Torras, 93b].
- La descripció dels dissenys en blocs en termes de mesures de distribució (parelles i quàdruples), i algunes propietats associades. En particular, la definició d'optimalitat en termes del nombre de quàdruples i d'uns.
- La definició d'un repertori de funcions de cost per a la generació de dissenys en blocs.
- La definició de les funcions de tipus  $H$ , basades en la minimització del l'error quadràtic dels camps locals [Bofill & Torras, 93a].
- La definició d'una nova estratègia per a les xarxes optimitzadores, basada en la cerca cooperativa [Bofill et al., 95].
- L'ús dels dissenys en blocs com a banc de proves per a l'avaluació experimental i comparació entre algorismes d'optimització.
- La definició d'una nova classe de configuracions combinatories, els dissenys màximament equilibrats, que elimina les restriccions d'admissibilitat sobre els paràmetres, i algunes de les seves propietats.



- Finalment, la generació experimental de dissenys màximament equilibrats.

Relacionats amb aquest treball, però no inclosos en la memòria perquè s'aparten de la línia argumental, s'han realitzat també els estudis que s'esmenten a continuació. En les simulacions preliminars, i sense una verificació sistemàtica, els resultats obtinguts amb la màquina de Boltzmann determinista (apartat 1.2.2) no eren superiors als de la recuita simulada. Aquesta va ser una raó de més per excloure-la del repertori. Un altre estudi realitzat va ser l'ús d'heurístiques senzilles per a reduir el nombre de configuracions isomorfes<sup>1</sup>. La reducció era important, però la seva aplicació era més apta per a estratègies de cerca enumeratives. Es va analitzar també l'aplicabilitat dels DMEs a la generació de codis binaris de màxima distància mínima, però les configuracions obtingudes eren massa petites, i no permetien millorar cap de les cotes conegudes. Finalment, es va intentar l'aplicació dels DMEs per a la generació de matrius d'interconnexió en sistemes multibus [Torràs & Bofill, 89], però l'enfoc no permetia satisfer totes les propietats desitjades.

## 7.2 Resultats de l'avaluació experimental

### 7.2.1 Aspectes metodològics

Pel que fa a la metodologia experimental, i de manera esquemàtica, les observacions que segueixen fan referència a les dificultats que es presenten a l'hora de *comparar algorismes* [Bofill & Torràs, 96]. En primer lloc, algunes consideracions sobre el context:

- La generació de dissenys és un problema NP-complet. No és NP-hard perquè podem determinar amb facilitat si una determinada configuració correspon a una solució.
- La generació de dissenys és un *problema de decisió* que, amb algorismes no deterministes, es descriu com un procés de Bernouilly amb probabilitat d'èxit  $p$ .
- En la generació de dissenys, no té sentit parlar de "solucions aproximades".
- En l'avaluació experimental, cal tenir en compte els *factors* que intervenen i, quan l'espai és molt gran, és necessari recórrer a processos independents d'entrenament i prova, basats en l'optimització experimental i la generalització.
- Alguns algorismes utilitzen més recursos que d'altres. Per tant, el resultat experimental ha de tenir en compte el cost, i cal treballar amb unitats insensibles a la implementació (cost  $z$  de *relaxació* fins la primera solució, en iteracions).

---

<sup>1</sup>La detecció d'isomorfisme ja és de per si un problema NP-hard.

- Si cal fer una *estimació* del resultat (resposta  $w$  de l'experiment paral·lel), el seu comportament estadístic és determinant per a l'aplicació de les tècniques d'anàlisi quantitatives.
- El coneixement teòric del procés ajuda a l'experimentació.

Les incidències observades, a la pràctica, són les següents:

- La formulació dels factors amaga *decisions implícites*. Aquest ha estat el cas en el repertori de problemes ( $u/vb \leq 0.5$ ), en l'ordre d'actualització, o en el tractament dels camps locals nuls.
- La variable resposta  $w$  adoleix de tres “defectes” importants. Possiblement no és normal, la seva variança és variable, i inclou casos de fracàs complet, que invaliden les tècniques de comparació objectives.
- La dispersió de la mesura augmenta quan els resultats de la cerca són pobres (que és quan més precisió cal), i dificulta l'optimització de paràmetres.
- Les interaccions entre factors són fortes, i amaguen els efectes principals. Això s'il·lustra, especialment, en l'elecció de la composició d' $F_{uthvq}$ .
- L'estandardització dels paràmetres, en el cas de CC, no és gens clara.

Per tot això, les comparacions són imprecises i la generalització de les decisions és necessàriament pobre. Tot i que els resultats finals han sigut clars, sempre queden alguns dubtes:

- Són fiables les decisions que s'han pres?
- Les alternatives no explorades, són realment més dolentes?
- Fins on es poden extrapolar les conclusions?

En conclusió, l'avaluació experimental d'algorismes és de per si *un problema difícil*.

## 7.2.2 Conclusions dels experiments

També de manera esquemàtica, i a partir dels resultats experimentals, les conclusions obtingudes són les següents:

- Resultats generals:
  - En l'espai experimental de prova, la combinació de la recuita simulada (amb paràmetres estàndard) amb la funció de cost  $F_{uq}$  ha estat, clarament, el millor *algorisme* de cerca.

- El nombre de problemes resolts decreix a mesura que augmenta el seu tamany. De 70 problemes amb  $500 < vb \leq 1000$ , només s’han pogut resoldre 8 dissenys, el més gran d’ells amb  $vb = 961$ .
- Les estratègies de cerca:
  - Malgrat l’elevat cost d’un descens, la eficàcia d’RS ha estat tan alta que ha conduït als resultats més eficients. Només un dels 25 problemes de prova ha quedat irresolt. El coneixement del procés ha estat útil per ajustar els paràmetres estàndard, i el comportament experimental d’RS ha estat molt “amical”: bons resultats i, en conseqüència, desviacions petites (bona precisió), pocs casos de fracàs complet, i una capacitat de generalització molt bona, tant respecte als problemes com a les funcions de cost. Per a RS, tots els mínims apareixien en la fase estocàstica de la cerca.
  - La millor qualitat de CD ha estat el baix cost d’un descens, conduïnt a un eficiència acceptable per als problemes que ha estat capaç de resoldre (13 dels 25). Tot i així, tal com era d’esperar, la seva eficàcia ha estat la més baixa. Malgrat tot, CD ha resultat molt útil com a estratègia de referència.
  - La ceca cooperativa, l’estratègia nova, ha estat bona en termes del nombre de casos resolts (22 dels 25). Tot i estar per darrera d’RS, ha superat netament a CD en termes d’eficàcia, però donat que és força costosa, en tots els problemes resolts per les dues la seva eficiència ha estat inferior. Els resultats amb paràmetres òptims, però, són força prometedors, si trobessim la manera d’estandarditzar-los. La manca de coneixement del procés no ha estat d’ajuda, i el seu comportament experimental ha estat incòmode: nombre de fracassos complets elevat, desviacions altes (menys precisió), baixa sensibilitat als paràmetres, i una generalització dels paràmetres molt pobre. La variant que ha resultat millor, cas a cas, és CCfd. Per a CC, els mínims no apareixen a la fase cooperativa de la cerca, sino en la fase descendent subseqüent.
- Les funcions de cost:
  - En l’espai experimental de prova, la millor funció de cost per a *totes* les estratègies ha estat la més senzilla,  $F_{uq}$ , amb una composició fixa. De fet, és la funció que s’havia fet servir en l’entrenament de les estratègies i, per tant, la més ben adaptada.
  - La segona millor funció de cost ha estat  $F_{uthvq}$ , la més completa. Tot i així, cal recordar que la seva composició havia estat més ben avaluada que la resta.  $F_{uthvq}$  és l’única estructura funcional apta per a qualsevol DME.

- Durant l'etapa d'entrenament, el factor funció de cost ha demostrat una interacció molt forta amb el factor problema. Com a conseqüència del procés de selecció utilitzat, la generalització de les funcions de cost finalistes no ha quedat completament verificada, especialment en relació a les altres dues estratègies (RS i CC).
- Els resultats amb les funcions de tipus  $H$  han estat pobres, com es descriu en treballs anteriors.
- Els problemes:
  - S'ha pogut observar que la dificultat dels problemes té una component intrínseca, tot i que la interacció entre els tres factors ha estat alta.
  - El cost de relaxació d'un descens mostra una lleu tendència a créixer amb el tamany del problema, amb diferents comportaments en funció de les estratègies i en funció dels resultats del descens (èxit/fracàs).
  - En línies generals, a mesura que els problemes es fan grans les dificultats de resolució augmenten, tant pel que fa al nombre d'intents com al cost fins la primera solució.
  - Per a problemes de tamany similar, els DMEs han estat lleugerament més difícils de resoldre, però no fan servir la millor funció de cost. Els casos amb  $v \leq b$  han resultat més fàcils.
- La variable resultat:
  - La variable resultat, el cost  $z$  fins la primera solució, mostra una distribució qualitativament semblant a la geomètrica, amb desviacions lleugeraments més petites que els valors esperats.
  - La distribució de  $z$  ve dominada pel nombre de descensos  $y$  fins la primera solució  $i$ , amb una desviació més petita, el cost  $c$  d'un sol descens actua principalment com un factor d'escala.
  - L'estimació d' $E(z)$  per mitjà de la resposta  $w$  de l'experiment paral·lel sembla ser prou acurada. Però el seu comportament estadístic adoleix dels inconvenients mencionats anteriorment.

### 7.3 Discussió

Com a problema a optimitzar, la generació de dissenys en blocs ha estat un banc de proves interessant, i els resultats del nostre millor algorisme sobre els problemes del conjunt de prova han estat bons. La comparació amb altres tècniques de generació de dissenys és difícil ja que, com hem mencionat a l'apartat 1.1, algunes famílies (infinites) de dissenys es poden construir analíticament, mentre que els treballs descrits

a la literatura referents a tècniques computacionals solen estar centrats a la resolució de subfamílies o casos oberts concrets. En particular, no hem trobat cap recull sistemàtic de resultats en funció del tamany del problema. El cas d'existència no resolt més petit actualment correspon al problema d94 [McKay & Radziszowski, 96], amb  $vb = 726$ . Això significa que, avui per avui, les tècniques computacionals no permeten l'anàlisi exhaustiva de problemes d'aquest tamany. Al capdevall, amb la nostra estatègia heurística hem pogut solucionar-ne dos de més grans (d99 i d123). L'ús dels disseny en blocs per comparar funcions i estratègies ha estat productiu, i el problema ha demostrat ser prou difícil per discriminar la qualitat dels algorismes.

L'habilitat dels models neuronals per trobar "bons" mínims no ha estat útil en el nostre cas, ja que els *pseudo*-òptims no serveixen (la generació de dissenys és un problema de decisió). La nostra anàlisi de l'optimalitat d'una configuració en termes de la distribució dels uns descriu algunes propietats dels dissenys en blocs. Aquest coneixement del problema, si bé ha estat a vegades necessari en la definició dels algorismes (estructura i òptim de les funcions de cost, funcions de tipus  $H$ , normalització de la temperatura computacional en RS), no s'ha utilitzat explícitament per a guiar les estratègies de cerca. Tot i així, un coneixement més profund seria útil per caracteritzar els resultats experimentals (dificultat dels problemes, coeficients de composició i paràmetres òptims de les estratègies), i per dissenyar nous algorismes de propòsit específic. En concret, la reducció d'isomorfismes podria resultar molt útil per reduir l'espai de cerca i evitar la competència entre solucions equivalents.

L'aportació dels DMEs com a configuració combinatòria queda pendent de validació. La seva formulació és atractiva, perquè és compacta i senzilla, però sobretot perquè elimina les restriccions d'admissibilitat tot oferint una definició precisa d'òptim. Malgrat tot, la seva aplicabilitat a l'anàlisi d'experiments s'ha de verificar, així com la seva optimalitat en termes d'anàlisi combinatòria [Kiefer, 80]. De fet, aquest tipus d'optimalitat combinatòria es podria provar d'utilitzar com a font de funcions de cost alternatives.

Mentre el problema defineix l'espai de cerca, la funció de cost defineix el paisatge, i la clau d'un algorisme eficaç és l'adequació de l'estratègia al darrer. Per a algorismes basats en la cerca local descendent, com els nostres, la propietat del paisatge que fa millorar els resultats és una relació òptims a mínims locals ben alta. En aquest treball, l'estudi s'ha realitzat en termes d'èxit/fracàs, però també seria útil analitzar altres aspectes del paisatge tals com el tamany i topologia de les conques d'atracció, les propietats i estructura dels mínims locals, la seva relació amb els òptims, la topologia del recorregut de la cerca, etc.

La nostra condició necessària i suficient d'optimalitat (teorema 2.4), és la definició de disseny més breu possible en termes de les variables d'estat, i condueix a la funció  $F_{uq}$ . Aquesta funció, la que millors resultats experimentals ha obtingut, és una prova de la capacitat descriptiva de les quàdruples. Tot i així, donat que és més fàcil distribuïr les quàdruples quan el nombre d'uns és baix, els seus inconvenients principals són les baixes prestacions per a problemes amb molts uns (com en el cas dels complementaris), i l'esbiaixament dels mínims locals cap a configuracions amb

$o < u$ . Això no és necessàriament un inconvenient, si aquest comportament lateral ja ens agrada i, de fet, l'elecció d'una funció de cost o altra, que porta a mínims locals amb diferents propietats, ens dona l'opció d'escollir els *pseudo*-òptims. La funció  $F_{q\bar{q}}$ , hauria d'equilibrar els espais original i complementari però, malhauradament, els dos subobjectius en aquest cas són difícils, i els resultats, com hem vist, han estat pobres. La segona millor funció,  $F_{uthvq}$ , apta també per als DMEs, suggereix que la redundància de subobjectius no és necessàriament dolenta, però té l'inconvenient que cal ajustar experimentalment els coeficients de composició. Forçats per la necessitat d'espurgar opcions, la selecció de funcions s'ha fet únicament en base a l'estratègia CD, però caldria contrastar les altres combinacions per assegurar que no hi ha alternatives més bones.

Les funcions de tipus  $H$  eren atractives en la seva formulació, fent ús del coneixement dels òptims però, a part de la seva complexitat computacional, el seu principal inconvenient és que fan créixer fortament el nombre de mínims locals. Precisament per això, però, podrien resultar molt útils per a estratègies enumeratives basades en la purga d'alternatives.

Pel que fa a les estratègies, moltes alternatives han quedat per explorar. Per a la RS l'única regla de decisió que hem provat és la de Metropolis. La funció sigmoïdal és l'alternativa natural. Però, en relació a l'esbiaixament de les quàdruples cap als zeros, una possible millora seria assignar els uns de manera determinista i els zeros de manera estocàstica. Amb un rang de temperatures estret (per tal de minimitzar el cost), una llei de decrement lineal permetria tal vegada millorar la tolerància en l'elecció de la temperatura crítica. La granularitat temporal petita i la llei de decrement suau han resultat prou eficients. L'extensió natural de la recuita simulada és la màquina de Boltzmann determinista, i els models derivats basats en la teoria del camp mitjà. Tot i que probablement no aconseguiríem reduir el cost, donat que aquests models extenen l'espai de cerca a l'interior de l'hipercub binari, podrien potser millorar l'eficàcia de la cerca. Tal com es proposa a [Van den Berg, 96], el model generalitzat de Hopfield permet formes d'exploració alternatives en que objectius i restriccions es satisfan simultàniament.

L'estratègia de la cerca cooperativa no ha estat prou eficaç per resultar competitiva en termes d'eficiència, però ha resolt un nombre elevat de problemes. Els paràmetres han estat ajustats per optimitzar  $z$ . Optimitzant  $y$  potser augmentaria aquest nombre, i obtindriem esquadres més grans. Amb  $S = 2$ , el valor estàndard que s'ha utilitzat, el nombre de mostres és massa petit per poder interpretar el comportament de l'esquadra en termes de camp mitjà. Donat que tots els mínims apareixen més tard, la verdadera contribució de la fase cooperativa és la cerca d'un bon estat inicial per a la subseqüent fase descendent. La topologia focalitzada ha obtingut els millors resultats, i les pitjors prestacions de la topologia dispersa podrien ser conseqüència de la competència entre solucions isomorfes entre si. Donat que els paràmetres estàndard obtinguts recorren tot el ventall de radis, la hipòtesi d'un radi crític no sembla molt versànt. El mecanisme de contracció, com hem vist, també fa cerca descendent, competint en rellevància amb la pròpia relaxació

de l'esquadra. En conjunt, però, cal aprofundir en l'anàlisi de la CC, per poder-la ajustar més bé. Com abans, moltes alternatives han quedat per explorar. Podriem començar amb esquadres grans, i anar-les descomposant progressivament. O podriem escollir les unitats variables d'una subregió de la matriu d'incidència, a fi de concentrar la cerca cap a una única solució.

La cerca cooperativa, donat que és una estratègia basada en una població (els membres de l'esquadra), que busca estats inicials bons, té alguns punts en comú amb alguns treballs de formulació recent, com ara [Boyan & Moore, 97, Baluja, 96, De Bonet et al., 96]. Les conclusions d'aquests i d'altres treballs similars també s'haurien de tenir en compte per provar de millorar les prestacions de CC.

La hipòtesi de que els òptims es troben en regions de baixa energia mitjana sembla en conjunt confirmada, però la gran eficiència d'RS, juntament amb la millora de CC quan les esquadres són petites, semblen indicar que les solucions es troben més en els *camins* que no pas en les *regions* d'energia mitjana decreixent.

Donada la complexitat del problema, és impossible trobar una solució teòrica pel cas general, i l'avaluació experimental es fa inevitable. La metodologia que hem fet servir ha resultat útil per reduir l'espai d'experimentació a dimensions abastables. Però, precisament per això, moltes alternatives han quedat inexplorades i no es poden extrapolar els resultats. En concret, l'avaluació de les funcions de cost podria ser més sistemàtica. En la definició d'*y* i *z*, l'ús d'estadístiques truncades per tractar els casos de fracàs complet permetria fer comparacions quantitatives, que ajudarien a identificar més aviat els factors experimentals més rellevants. L'ús del cost esperat fins la primera solució com a variable resultat ha estat prou útil per comparar algorismes sobre un conjunt de problemes *específic*. Però, si es tracta d'avaluar la seva escalabilitat cap a problemes cada vegada més grans, el vertader indicador hauria de ser, provablement, el tamany crític a partir del qual el nombre de problemes resolts decau considerablement. La transferència bidireccional de tècniques d'optimització entre el domini del model i el domini experimental és una altra idea a tenir en compte.

El treball que hem presentat aquí descriu la cerca dins de la cerca, en quatre nivells concèntrics: En el domini objecte, el *problema* (un disseny) satisfà certes propietats d'optimalitat de la teoria combinatòria. A continuació, en el nivell algorísmic, els *models* es defineixen per a buscar solucions en l'espai de configuracions del problema. Aleshores, a nivell *experimental*, busquem el millor algorisme que, dins de l'espai dels models, té millors prestacions sobre el problema. Finalment, i per sobre de tots, la *recerca* explora els tres nivells anteriors.

La *dificultat* del problema s'extén a tots els nivells, i delimitar alternatives en un pla no és més, moltes vegades, que transferir les dificultats a un altre. La construcció analítica de dissenys només és possible per a algunes famílies concretes, les tècniques computacionals s'encallen a partir d'un cert tamany, i l'avaluació experimental no pot abastar tots els models possibles. La recerca, però, ho ha de seguir intentant, amb l'esperança, si no de resoldre el problema, almenys de millorar els resultats. La classe NP, en definitiva, podria *no* ser P.

## 7.4 Perspectives

Recopilant la discussió anterior, i recorrent una mica al pragmatisme, les perspectives de continuació d'aquest treball que proposem són les següents:

- Incorporació del coneixement del problema a les estratègies de cerca, a fi de millorar la generació de dissenys. En particular, tractament de l'isomorfisme per focalitzar la cerca en una única solució.
- Aplicabilitat dels DMEs a l'anàlisi d'experiments, i anàlisi de les seves propietats d'optimalitat combinatòria. Possibilitats d'ús de les darreres com a funció de cost.
- Aprofundiment en l'anàlisi de les propietats dels DMEs.
- Utilització de les funcions de cost proposades amb altres estratègies de cerca. En particular, ús de les funcions de tipus  $H$  amb estratègies enumeratives.
- Anàlisi del paisatge de cerca, i aplicació a les estratègies.
- Utilització del coneixement del problema per caracteritzar la dificultat dels problemes, i per ajustar els paràmetres de les estratègies.
- Caracterització i millora de les estratègies proposades, en termes de la discussió precedent, i en relació a tècniques similars de la literatura. Noves propostes.
- Refinament de les tècniques d'anàlisi experimental emprades.
- Ús dels dissenys en blocs com a banc de proves per a la comparació de tècniques d'optimització. Extensió a d'altres algorismes, tant neuronals com no neuronals.



# Agraïments<sup>1</sup>

Tot al llarg d'aquesta etapa, sou molts els que m'heu ajudat i fet costat en la realització d'aquest treball. Voldria mencionar en primer lloc a la Carme, pel seu treball de direcció i el seu recolzament personal, i a l'Enric F., que amb el seu treball fi de carrera em va donar la darrera empenta. També vull mencionar a la Maite, la Mònica i la Mireia, que ho han viscut des de casa, i a la família propera, que també els ha tocat patir.

Molts de vosaltres, no us puc nombrar a tots, m'heu atès en les meves consultes. L'Enric M., amic i consultor particular en tots els àmbits, el grup de combinatòria (l'Oriol i l'Anna), els estadístics (en especial l'Eulàlia, la Lupe, la Karina, la Marta, en Manel i en Paco), els "sistems" (el plat fort se'l va endur l'Uri), i diversos (en Xavi, el Paco, la Isabel). Cito també als meus primers companys en l'àmbit de les xarxes neuronals (en Josep Maria, el Pepe, l'Enric, la Gabriela, en Gordon, en Vicente i la resta).

En el transcurs del dia a dia, en Leandro és qui ha patit primer les conseqüències, i l'Eugeni i en Josep Maria els qui van saber-ho arreglar. Però agraeixo també el suport i companyia de la resta de companys de feina, i l'ajut insubstituïble del personal d'administració i serveis. També vull mencionar alguns companys de penúries (l'Andreu, en Jordi, la Patricia, en Toni), i penúries (en Joan, en Sergi, en Jaume, la Magda), i animar des d'aquí als que encara els falta un tram. Finalment, en l'àmbit personal, als col·legues que m'esperen per poder-ho celebrar (en Joan, la Marta, en Ramón, la Lulu, la Joana, en Ferran, en Juan Carlos, en Jordi, en Lluís, l'Ana, els bufis i companyia, els de "La Pedra", els coralinos, ...) incloent-hi, per extensió, les col·les respectives. El meu agraïment, en definitiva, per a tots vosaltres que m'esteu llegint.

Barcelona, Primavera de 1997.

---

<sup>1</sup>Aquest treball ha estat parcialment finançat pel projecte "Redes neuronales realimentadas y autoorganizativas para la identificación, el control y la optimización de sistemas", CICYT TIC-91-0423.



# Bibliografia

- [Aarts & Korst, 87] Aarts E.H.L, Korst J.H.M., "Boltzmann machines and their applications", *Proc. PARLE*. Springer-Verlag. Lecture Notes in Computer Science, Vol 258, p 34-50, 1987.
- [Aarts & Korst, 88] Aarts E.H.L, Korst J.H.M., *Simulated Annealing and Boltzmann Machines*, Wiley Interscience, 1988.
- [Ackley et al., 85] Ackley D.H., Hinton G.E.; Sejnowsky T.J., "A Learning Algorithm for Boltzmann Machines", *Cognitive Science*, Vol 9, 147, 1985.
- [Anderson, 89] Anderson, Ion, "A first course in combinatorial mathematics", *Oxford Applied Mathematics and Computer Science Series*, Secona edició, 1989.
- [Arbib, 89] Arbib M., *The metaphorical brain 2: Neural Networks and beyond*, John Wiley and Sons, 1989.
- [Baluja, 96] Baluja S., "Genethic Algorithms and Explicit Search Statistics", *Advances in Neural Information Processing Sytems*, (pendent de publicació).
- [Barto et al., 81] Barto A.G., Sutton R.S, Brouwer P.S., "Associative search network: a reinforcement learning associative memory", *Biological Cybernetics*, Vol 40, pp 201-211, 1981.
- [Bofill, 92] Bofill P. "Xarxes Neuronals Artificials", *Butll. Soc. Cat. Ciènc.*, Vol. XIV, Núm. 2, pp 247-260, 1994. Presentat a la *Universitat Catalana d'estiu de Prada*, estiu de 1992.
- [Bofill et al., 95] Bofill P., Fontdecaba E., Torras C., "Optimization Networks for the Generation of Block Designs", *Journal of Artificial Neural Networks*, Vol 2(4), pp 302-312, 1995.
- [Bofill et al., 90] Bofill P., Manyer J., Millán J.R., Salvans V., "A systolic algorithm for back-propagation: mapping onto a transputer network", J.Sender, T.Addis, Eds., *Symbols vs. Neurons*, IOS Press, p 32-46, 1990.
- [Bofill et al., 89] Bofill P., Millán J.R., Torras C., "Short-term and long-term optimization in neural networks. Two applications", *Revista de Matemáticas Aplicadas*, Vol 12, pp 63-72, Universidad de Chile, 1991.

- [Bofill & Torras, 96] Bofill P., Torras C., "Comparing Optimizing Neural Networks on the Generation of Block-Designs: Methodological issues", ponencia oral al "Workshop on Nature-Inspired Algorithms for Combinatorial Optimization", Jagota, A., (chair), Snowmass, CO, Desembre 96.
- [Bofill & Torras, 93a] Bofill P., Torras C., "A Second-Order Optimization Strategy for the Generation of Block Designs", *Proc. WCNN'93, Portland (Oregon)*, INNS-Press, Vol 4, pp 323-327, Juliol 1993.
- [Bofill & Torras, 93b] Bofill P., Torras C., "Higher-Order Networks for the Optimization of Block Designs", Mira, Cabestany, Prieto, Eds. *New Trends in Neural Computation*, Springer-Verlag, Lecture Notes in Computer Science, Vol 686, pp 114-118, Juny 1993.
- [Boyan & Moore, 97] Boyan J.A., Moore A.W., "Using Prediction to Improve Combinatorial Optimization Search", *AISTATS-97*, (pendent de publicació).
- [TMC] *CM-200 Technical Summary*. Thinking Machines Corporation. June 1991.
- [Corneil & Mathon, 78] Corneil D.G., Mathon R.A., "Algorithmic Techniques for the Generation and Analysis of Strongly Regular Graphs and Other Combinatorial Configurations", *Ann. of Discrete Mathematics*, North Holland Publishing Company, Vol. 2, pp. 1-32, 1978.
- [De Bonet et al., 96] De Bonet, J., Isbell, C., Viola, P., "MIMIC: Finding Optima by Estimating Probability Densities", *Advances in Neural Information Processing Systems*, (pendent de publicació).
- [Fisher, 40] Fisher R.A., "An examination of the different possible solutions of a problem in incomplete blocks", *Ann. Eugen.*, Vol 10, p 52-57, 1940.
- [Fontdecaba, 94] Fontdecaba F., "Algorismes per a la cerca de BIBD's", *Projecte Final de Carrera*, E.T.S.E. Telecomunicació de Barcelona. Dirigit per Pau Bofill. Setembre 1994.
- [Garey & Johnson, 79] Garey M.R. & Johnson D.S., *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, San Francisco, 1979.
- [Gibbons, 96] Gibbons P.B., "Computational Methods in Design Theory", *The CRC Handbook of Combinatorial Designs*, pp 3-41, 1996.
- [Gibbons et al., 77] Gibbons P.B., Mathon R.A., Corneli D.G., "Computing Techniques for the Construction and Analysis of Block Designs", *Utilitas Mathematica*, Vol 11, pp 161-192, 1977.

- [Goles & Matamala, 94] Goles E., Matamala M., "Dynamical and Complexity Results for High Order Neural Networks", *International Journal of Neural Systems*, Vol. 5(3), pp 241-252, 1994.
- [Grossberg, 87] Grossberg S., "Competitive learning: form interactive activation to adaptive resonance", *Cogn. Science*, Vol 11, N 1, p 23-63, 1987.
- [Gutzmann, 87] Gutzmann K.M., "Combinatorial Optimization Using a Continuous State Boltzmann Machine" *Proceedings of IANN Conference*, San Diego, 1987.
- [Hall, 86] Hall M., *Combinatorial Theory*, Ed. John Wiley & Sons, Second Edition 1986.
- [Hertz et al., 93] Hertz J., Krogh A., Palmer R.G., *Introduction to the Theory of Neural Computation*, Ed., Addison-Wesley, Gener 1993.
- [Hinton et al., 84] Hinton G.E., Sejnowsky T.J., Ackley D.H., "Boltzmann machines: Constraint satisfaction networks that learn", *Technical Report CMU-CS-84-119*, Carnegie-Mellon Univ., 1984.
- [Hopfield, 82] Hopfield J.J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proc. Nat. Academ. Sciences USA*, Vol 79, 2554-2558, 1982.
- [Hopfield, 84] Hofield J.J., "Neurons with graded response have collective computational properties like those of two-state neurons", *Proc. of the National Academy of Sciences*, USA 81, pp 3088-92, May 1984.
- [Hopfield & Tank, 85] Hopfield J.J., Tank D.W., "'Neural' Computation of Decisions for Optimization Problems", *Byological Cybern.*, Vol. 52, pp. 141-152, 1985.
- [Jagota & Regan, 97] Jagota A.K., Regan K.W., "Performance of Neural Net Heuristics for Maximum Clique on Diverse Highly Compressible Graphs", *Journal of Global Optimization*, Vol. 7, pp. 1-26, 1997.
- [Jagota & Wang, 96] Jagota A., Wang X., "Cycle-Free Dynamics of a Cluster-Competitive Net", *Proceedings of WCNN*, invited talk, San Diego, 1996.
- [Jordan, 86] Jordan M.J., "Attractor Dynamics and Parallelism in a Connectionist Sequential Machine", *VIII<sup>th</sup> A. Meet. Cogn. Scie. Soc.*, Hillsdale, 1986.
- [Joya et al., 91] Joya G., Atencia M.A., Sandoval F., "Application of High-Order Hopfield Neural Networks to the Solution of Diophantine Equations", *Proc. IWANN-91*, pp. 395-400, 1991.
- [Kiefer, 80] Kiefer J., "Optimal Design Theory in relation to Combinatorial Design", *Annals of Discrete Mathematics*, Vol. 6, pp. 225-241, 1980.

- [Kirkpatrick et al., 83] Kirkpatrick S., Gelart Jr., Vecchi M.P. "Optimization by simulated annealing", *Science*, Vol 220, N 671, 1983.
- [Kohonen, 87] Kohonen T., *Self organization and associative memory*, Springer-Verlag, 1987.
- [Koiran, 94] Koiran P., "Dynamics of Discrete Time, Continuous State Hopfield Networks", *Neural Computation*, Vol. 6, pp. 459-468, 1994.
- [LeCun, 85] LeCun Y., "A learning scheme for asymmetric threshold networks", *Proc. of Cognitiva 85*, Paris, 1985.
- [Lippmann, 87] Lippmann R.P., "An introduction to computing with neural nets", *IEEE ASSP magazin*, April-87, 4-22, 1987.
- [López et al., 89] López F.J., Acevedo M.J., Jaramillo M.A., "Redes neuronales. Evolución histórica", *Mundo electrónico*, N. 197, pp. 57-71, Juliol 1989.
- [Mathon & Rosa, 90] Mathon R., Rosa A., "Tables of parameters of BIBD with  $r \leq 41$  including existence, enumeration and resolvability results: an update", *Ars Combinatoria*, Vol 30, December, Winnipeg, Canada, 1990.
- [Mathon & Rosa, 96] Mathon R., Rosa A., " $2-(v, k, \lambda)$  Designs of Small Order", *The CRC Handbook of Combinatorial Designs*, pp 3-41, 1996.
- [McCulloch & Pitts, 43] McCulloch W.S., Pitts W., "A Logical Calculus of the Ideas Imminent in Nervous Activity", *Bulletin of Mathematical Biophysics*, Vol. 5, pp. 115-133, 1943.
- [McKay & Radziszowski, 96] McKay B.D., Radziszowski S.P., "Towards Deciding the Existence of  $2-(22, 8, 4)$  Designs", *Journal of Combinatorial Mathematics and Combinatorial Computing*, Vol. 22, pp. 211-22, 1996.
- [Metropolis et al., 53] Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E., "Equation of State Calculations for Fast Computing Machines", *Journal of Chemical Physics*, Vol 21, pp 1087-1092. 1953.
- [Millán & Bofill, 89a] Millán J.R., Bofill P., "Learning by back-propagation: Computing in a systolic way", *Proc. of PARLE-89*, Lecture Notes in Computers Science, 366, p 235-254, Springer-Verlag, 1989.
- [Millán & Bofill, 89b] Millán J.R., Bofill P., "Learning by back-propagation: A systolic algorithm and its transputer implementation", *Neural Networks World*, Vol 1, N 3, p 119-137, July, 1989.
- [Montgomery, 91] *Design and Analysis of Experiments*, John Wiley & sons, tercera edició, 1991.

- [Peterson & Anderson, 87] Peterson C., Anderson J.R., "A Mean Field Theory Learning Algorithm for Neural Networks", *Complex Systems*, Vol 1, N 5, 995-1019, 1987.
- [Peterson & Södeberg, 89] Peterson C., Södeberg B., "A New Method for Mapping Optimization Problems onto Neural Networks", *Int. Journ. Neural Sys.*, Vol. 1, N. 1, pp. 3-22, 1989.
- [Rumelhart et al., 86] Rumelhart D.E., Hinton G.E., Williams R.J., "Learning representations by back-propagating errors", *Nature*, Vol 323, 533-536, 1986.
- [Rumelhart & McClelland, 86] Rumelhart D.E., McClelland J.L, (Eds.), *Parallel distributed processing, PDP*, Vols 1,2, MIT Press, 1986.
- [Sejnowski, 86] Sejnowski T.J., "Higher-Order Boltzmann Machines", *Proc AIP*, Snowbird 1986.
- [Street & Street, 87] Street A. P., Street D. J., *Combinatorics of Experimental Design*, Oxford Science Publications, Clarendon, Oxford 1987.
- [Szu, 86] Szu H., "Fast Simulated Annealing", *Proc. AIP Conference*, Snowbird, 1986.
- [Tagliarini & Page, 87] Tagliarini G.A, Page E.W., "Solving Constraint Satisfaction Problems with Neural Networks", *Proc. ICNN-87*, Vol 3, pp 741-747, 1987.
- [Torrás, 89] Torrás C., "Relaxation and neural learning: Points of convergence and divergence", *Jour. of Parallel Distributed Computing*, Vol 6, p 217-244, 1989.
- [Torrás & Bofill, 89] Torrás C., Bofill P., "A neural solution to finding optimal multibus interconnection networks", *Proc. IX Conf. de la Sociedad Chilena de Ciencias de la Computación*, Vol 1, p 446- 454, Juliol 1989.
- [Van den Berg, 96] Van den Berg J., "Neural Relaxation Dynamics, Mathematics and Physics of Recurrent Neural Networks with Applications in the Field of Combinatorial Optimization", *PhD thesis*, Erasmus University Rotterdam, 1996.
- [Yates, 35] Yates F., "Complex Experiments (with discussion)", *Journ. of the Royal Statist. Soc.*, Suppl. 2, 181-247, 1935.







